

Universidade do Minho
Escola de Engenharia

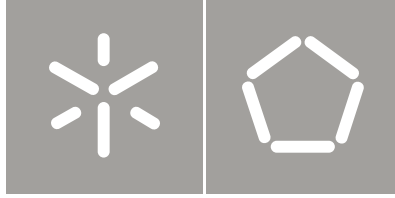
Merylin Ruth Cibrão dos Santos

Human-machine interface por olho humano

Merylin Ruth Cibrão dos Santos Human-machine interface por olho humano

UMinho | 2011

outubro de 2011



Universidade do Minho
Escola de Engenharia

Merylin Ruth Cibrão dos Santos

Human-machine interface por olho humano

Dissertação de Mestrado
Ciclo de Estudos Integrados Conducentes ao
Grau de Mestre em Engenharia Electrónica Industrial e Computadores

Trabalho efetuado sob a orientação do
Professor Doutor Fernando Ribeiro

AGRADECIMENTOS

Agradeço ao meu orientador, Dr. Fernando Ribeiro, por toda a ajuda e apoio prestados, essenciais para esta tese. Agradeço também a todo o pessoal do Laboratório de Automação e Robótica pelo material e espaço cedidos, tornando assim este trabalho menos inquietante.

De igual forma agradeço a todos os meus colegas de curso, em especial ao João Silva e Sérgio Silva, por todo o tempo dispendido e observações feitas.

Por último, um agradecimento especial à Margarita Oliveira e Pedro Guerreiro por todo o apoio moral, paciência e ânimo que me deram.

RESUMO

Nesta tese foi desenvolvida uma interface entre o humano e a máquina através do movimento do olho (*eye tracking*), de maneira a permitir que pessoas com tetraplegia usem o computador. Também no tema da tecnologia de assistência, foi desenvolvida uma interface gráfica com características especiais como, por exemplo, uma aplicação de teclado virtual para melhorar a comunicação do utilizador. Foi ainda desenhada uma estrutura de apoio capaz de fixar e manusear o sistema de forma simples e fácil.

O sistema foi desenvolvido tendo em mente o custo, que deve ser reduzido, e a sua funcionalidade, que deve trazer mais independência e autonomia ao utilizador.

Palavras-chave: seguimento do olho, tecnologia de assistência, interface humano-máquina.

ABSTRACT

In this thesis was developed an interface between human and machine through the movement of the eye (eye tracking), in order to allow people with tetraplegia to use the computer. Also on the topic of assistive technology, was developed a graphical interface with special features such as, for example, a virtual keyboard to improve user's communication. It was also designed a support structure that is able to attach and handle the system in a simple and easy way.

The system was developed keeping in mind the cost, which should be low, and its functionality, which should bring more independence and autonomy to the user.

Keywords: eye tracking, assistive technology, human-machine interface.

ÍNDICE

Agradecimentos	III
Resumo	V
Abstract	VII
Índice de figuras	XI
1. Introdução	1
1.1. Conceitos	2
1.1.1. <i>Bright pupil</i>	3
1.1.2. <i>Dark pupil</i>	4
1.2. Estado da Arte	5
1.2.1. <i>The Eyegaze Edge™, LC Technologies, Inc.</i>	5
1.2.2. <i>TM4, EyeTech Digital Systems</i>	6
1.2.3. <i>MyTobii P10, Tobii</i>	7
1.2.4. <i>EyeMax, DynaVox</i>	8
1.2.5. <i>IntelliGaze™, alea technologies</i>	9
1.2.6. <i>SeeTech®, Humanelektronik GmbH</i>	10
1.2.7. <i>ECO2 com ECOpoint, Prentke Romich Company</i>	11
1.2.8. <i>EyeWriter</i>	12
1.3. Tabela comparativa	13
2. Selecção de hardware e software	15
3. <i>Eye Tracking</i>	19
3.1. Detecção da pupila e do <i>glint</i>	19
3.1.1. Cor/luminosidade	20
3.1.2. Forma	20
3.2. Recolha e tratamento de dados	24

3.2.1.	Filtro de <i>Kalman</i>	24
3.3.	Mapeamento	27
3.3.1.	Calibração.....	27
3.3.2.	Transformação de coordenadas.....	28
3.4.	Fluxograma.....	31
4.	Interface.....	33
4.1.	Teclado	33
4.2.	Frases predefinidas.....	35
4.3.	Multimédia	35
4.4.	Internet browser	37
4.5.	Configurações.....	38
4.6.	Clique.....	40
5.	Estrutura de apoio	41
6.	Conclusão	43
6.1.	Características do sistema	44
6.2.	Propostas para trabalho futuro	44
7.	Referências.....	47
	Apêndice I.....	49
	Apêndice II.....	50

ÍNDICE DE FIGURAS

Figura 1. <i>Electro-OculoGraphy</i> . [3]	2
Figura 2. Posição da fonte de luz para <i>bright pupil</i> . [4]	3
Figura 3. Posição da fonte de luz para <i>dark pupil</i> . [4]	3
Figura 4. Método <i>bright pupil</i> . [5]	4
Figura 5. Método <i>dark pupil</i> . [5]	4
Figura 6. <i>The Eyegaze Edge Tablet #521</i> , <i>LC Technologies</i> . [6]	5
Figura 7. Teclado, <i>The Eyegaze edge</i> . [6]	6
Figura 8. <i>TM4</i> , <i>EyeTech Digital Systems</i> . [8]	7
Figura 9. <i>MyTobii P10</i> , <i>Tobii</i> . [9]	7
Figura 10. <i>EyeMax</i> e <i>Vmax</i> , <i>DynaVox</i> . [10]	8
Figura 11. Aplicações, <i>Vmax</i> . [10]	9
Figura 12. <i>IntelligazeTM</i> , <i>alea technologies</i> . [11]	10
Figura 13. <i>SeeTech®</i> , <i>Humanelektronik GmbH</i> . [12]	10
Figura 14. <i>ECO2</i> com <i>ECOpoint</i> , <i>Prentke Romich Company</i> . [13]	11
Figura 15. <i>EyeWriter</i> . [14]	12
Figura 16. Configuração do <i>EyeWriter</i> . [14]	12
Figura 17. <i>PlayStation® Eye</i> . [15]	15
Figura 18. Direita: 18 mm; esquerda: 6 mm.	16
Figura 19. Iluminador infravermelho de 12 led.	16
Figura 20. <i>Dark pupil</i> (esquerda) e <i>bright pupil</i> (direita).	19
Figura 21. Resultado do operador <i>canny</i>	21
Figura 22. Transformada de <i>Hough</i>	21
Figura 23. Gráfico da matriz de votos.	22
Figura 24. Transformada de <i>Hough</i> com raios diferentes.	22
Figura 25. Esquerda: olho quando se olha para algo claro. Direita: olho quando se olha para algo escuro.	23
Figura 26. Detecção do <i>glint</i>	23
Figura 27. Vector entre o <i>glint</i> e a pupila.	24
Figura 28. Resposta do filtro de kalman, com $Q = 1 \times 10^{-1}$ e $R = 0,1$	26
Figura 29. Resposta do filtro de kalman, com $Q = 1 \times 10^{-2}$ e $R = 0,1$	26
Figura 30. Resposta do filtro de kalman, com $Q = 1 \times 10^{-3}$ e $R = 0,1$	26
Figura 31. Resposta do filtro de kalman, com $Q = 1 \times 10^{-4}$ e $R = 0,1$	27
Figura 32. Pontos do ecrã (em cima). Pontos da pupila (em baixo).	27

Figura 33. Coordenadas da pupila.....	28
Figura 34. Rectas e quadrantes.....	28
Figura 35. Relação da pupila.....	29
Figura 36. Relação do ecrã.	29
Figura 37. Fluxograma de <i>eye tracking</i>	31
Figura 38. Menu.	33
Figura 39. Teclado.	34
Figura 40. Alerta.	34
Figura 41. Frases predefinidas.	35
Figura 42. Reprodutor de música.	36
Figura 43. Reprodutor de vídeo.	36
Figura 44. Visualização de imagens.....	36
Figura 45. Internet <i>browser</i>	37
Figura 46. Configuração.....	38
Figura 47. Pontos de calibração.	39
Figura 48. Estrutura de apoio.....	41
Figura 49. Mecanismo de selecção de alturas.	42
Figura 50. <i>Kernel</i> de rotulagem.....	49
Figura 51. Aplicação do <i>kernel</i>	49
Figura 52. Algoritmo de detecção de <i>blobs</i>	49

1. INTRODUÇÃO

Em 1963 foi publicado um software revolucionário que veio mudar a maneira como as pessoas interagem com os computadores. Desenvolvido por Ivan Edward Sutherland, a aplicação informática Sketchpad [1] é o antecessor do CAD (*Computer-aided design*) e foi o principal avanço para o GUI (interface gráfica do utilizador, do inglês *graphical user interface*). Desde então a interface entre o humano e a máquina tem sido explorada.

A *Association for Computing Machinery* (associação científica e educacional para maquinaria da computação) define esta interface como:

“Interacção humano-computador é a disciplina que se ocupa com o design, avaliação e implementação de sistemas computacionais interactivos para o uso humano e também com o estudo dos principais fenómenos que a rodeiam.” [2]

A naturalidade e facilidade com que o utilizador interage com a máquina são os aspectos mais importantes na interface, minimizando assim a barreira entre o utilizador e a máquina.

A motivação para a realização desta tese reside nas dificuldades encontradas por pessoas com deficiências motoras, onde a interacção com uma máquina seria vantajosa no dia-a-dia. Principalmente para as pessoas que não conseguem falar, a interacção através da conversação por escrita mostra-se importante por motivos de saúde mental e de inserção social.

O principal objectivo desta tese reside na realização de uma interface entre o humano e a máquina, onde o humano é tetraplégico (ou quadriplégico) sendo capaz de unicamente mover os olhos. Deve ser uma solução sem constrangimentos para o utilizador, isto é, sem dispositivos que se liguem ao humano, tais como óculos ou eléctrodos. Tem também de considerar as necessidades estruturais do utilizador através de uma estrutura de suporte e manuseamento do sistema.

Esta dissertação consiste então no desenvolvimento de um dispositivo de seguimento do olhar que permite ao utilizador interagir com a máquina, neste caso como se fosse um rato do computador. São também desenvolvidas aplicações com características especiais, dedicadas a este tipo de utilizador. E, finalmente, é apresentado o design de uma estrutura de apoio para todo o sistema.

1.1. CONCEITOS

Este capítulo tem como objectivo aprofundar os conceitos que vão ser referidos ao longo do relatório. *Eye tracking* ou *gaze tracking* são as expressões, do inglês, pelas quais globalmente é conhecida a acção de seguimento do olho ou do olhar, respectivamente. Isto pode então servir de interface para a utilização de variados serviços ou máquinas.

Para fazer o *eye tracking* podem ser usados vários métodos, desde *Electro-OculoGraphy* a métodos baseados em vídeo. No primeiro método, são ligados eléctrodos em redor dos olhos do utilizador, como se mostra na Figura 1, através dos quais se mede a diferença de potencial entre a pele. Este método não é contemplado pois é invasivo para o utilizador, o que vai contra os objectivos deste trabalho.



Figura 1. *Electro-OculoGraphy*. [3]

Os métodos baseados em vídeo são os mais usados para o *eye tracking* por não implicarem hardware tão preciso ou complexo. Para este caso pode-se recorrer a câmaras pouco dispendiosas e complementar com processamento de imagem. São destacadas duas vertentes deste método, *bright pupil* e *dark pupil*.

Tanto no *bright pupil* como no *dark pupil* são usadas fontes de luz infravermelha. No caso do *bright pupil*, a fonte de luz está alinhada com o eixo óptico da câmara, como descrito na Figura 2. No caso do *dark pupil*, a fonte de luz está afastada do eixo óptico da câmara, como descrito na Figura 3. Estes métodos são de seguida analisados.

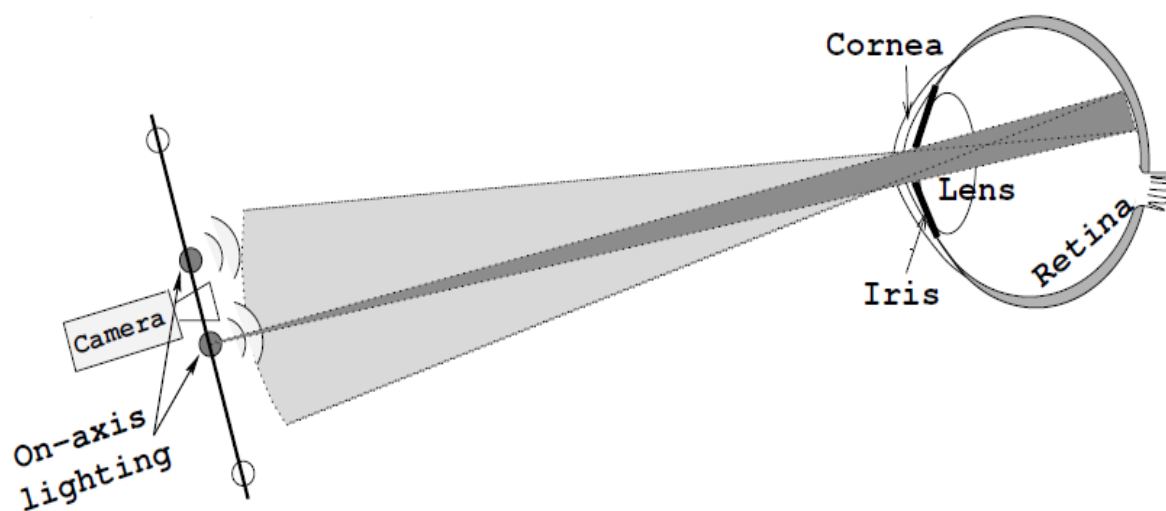


Figura 2. Posição da fonte de luz para *bright pupil*. [4]

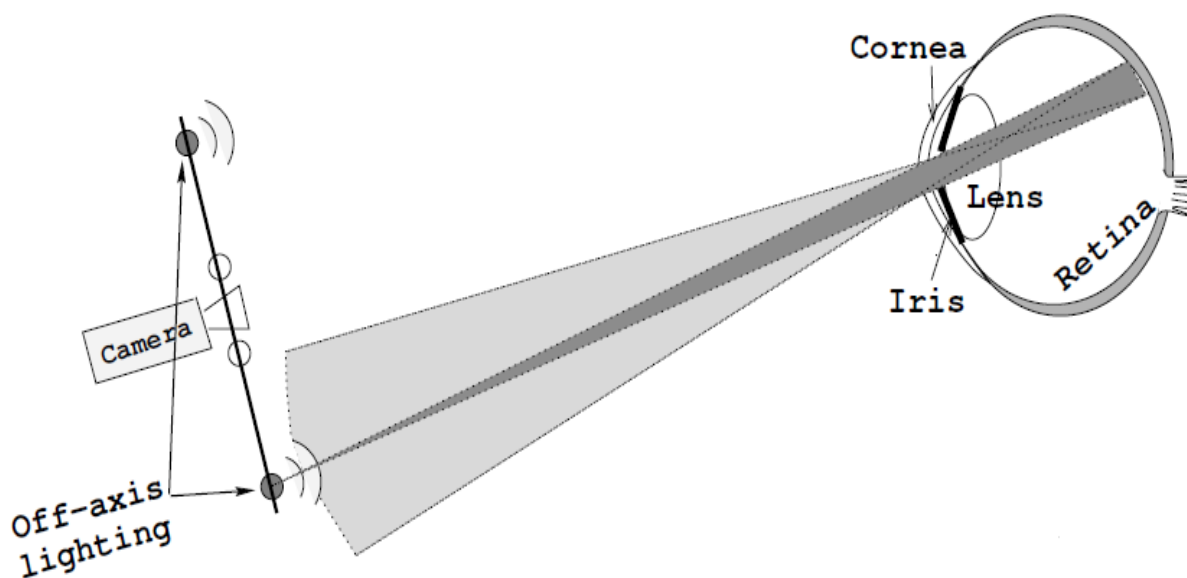


Figura 3. Posição da fonte de luz para *dark pupil*. [4]

1.1.1. **BRIGHT PUPIL**

No método *bright pupil*, devido à fonte de luz estar no eixo óptico da câmara, a reflexão dos raios infravermelhos ocorre, na sua maior parte, em direcção à câmara resultando numa pupila brilhante ou clara, como se pode ver na Figura 4. Este é o mesmo efeito presenciado nos olhos vermelhos numa fotografia aquando a utilização de um flash.

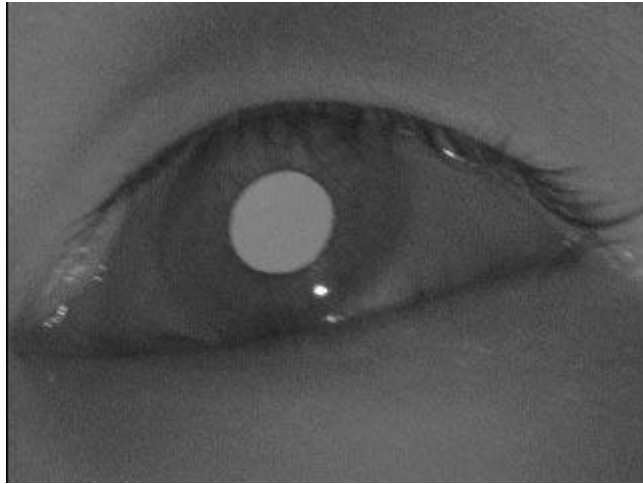


Figura 4. Método *bright pupil*. [5]

Deste método obtém-se uma pupila fácil de detectar, devido ao alto contraste. É mais eficaz quando usado em meios escuros, com pouca interferência de infravermelhos. Contudo é um método sensível ao tamanho da pupila e mostra ser mais exacto em olhos de íris clara, porque produzem maior contraste.

1.1.2. DARK PUPIL

Neste método, como a fonte de luz está afastada, a luz reflectida pelos olhos não coincide com o eixo da câmara, resultando numa pupila escura, como visível na Figura 5. Existe, no entanto, um ponto brilhante nesta figura, sendo este ponto chamado de *glint* (clarão, em português). O *glint* é o reflexo da fonte de luz infravermelha.

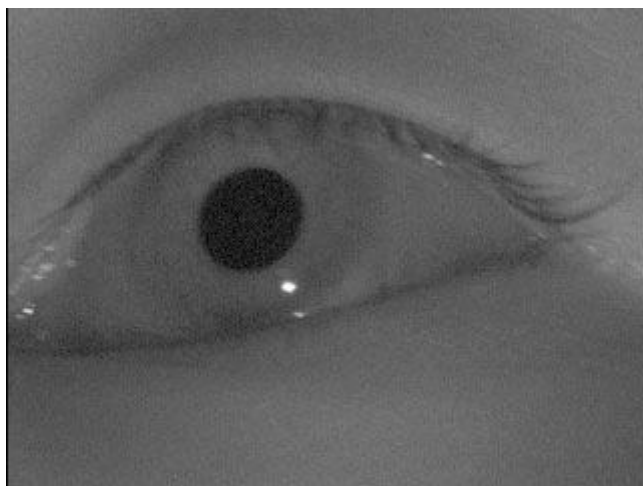


Figura 5. Método *dark pupil*. [5]

Este método, ao contrário do *bright pupil*, pode ser usado em meios com muita luz (como o exterior) e não é tão sensível ao tamanho da pupila. Porém, as pestanas e sombras podem causar falsos positivos na detecção da pupila.

1.2. ESTADO DA ARTE

Eye tracking aplicado a computadores para o movimento do rato é uma interface que ainda está em fase de maturação embora já existam produtos comerciais na área de tecnologia de assistência. Nesta secção serão abordadas estas alternativas, caracterizando-as de forma objectiva.

Estes sistemas são o ponto de partida para o desenvolvimento desta tese pois apresentam características que vão de encontro com os objectivos deste trabalho.

1.2.1. THE EYEGAZE EDGE™, LC TECHNOLOGIES, INC.

A *LC Technologies* apresenta uma solução de tecnologia de apoio para pessoas com deficiência física, o *The Eyegaze Edge™*. O utilizador apenas necessita de olhar para as teclas de controlo exibidas no monitor. Estão disponíveis 3 configurações diferentes deste sistema. Uma das configurações, apresentada na Figura 6, consiste numa solução leve (aproximadamente 2,200 kg), com um braço extensível podendo ser montado numa cadeira de rodas ou numa mesa.



Figura 6. *The Eyegaze Edge Tablet #521, LC Technologies.* [6]

O *The Eyegaze Edge Tablet* pode ser utilizado por pessoas desde os 3 anos e é ainda adaptável a uma gama ampla de variação de olhos como, por exemplo, a pálpebras caídas, pupilas assimétricas e olhos secos. O utilizador deve ter bom controlo de pelo menos um dos olhos. Pessoas com nistagmo (movimento repetitivo e involuntário de um ou dos dois olhos) ou estrabismo (perda do

paralelismo dos olhos) não serão capazes de fixar o olhar, fazendo com que o sistema não corresponda ao esperado.

Os softwares de interação com o utilizador incluídos nesta solução foram desenvolvidos tendo em conta as necessidades de pessoas com esclerose lateral amiotrófica, traumatismos cerebrais, paralisia cerebral, esclerose múltipla, atrofia muscular, entre outros. Por exemplo, a síntese de fala é importante para pessoas que querem manter uma conversa fluida. O programa “Teclado”, visível na Figura 7, foi usado na escrita de 4 livros. Outros exemplos são o “Frases”, “Aprendizagem” e “Configurações do programa”.

ESC	<-	->	^	v	Hom	End	PgU	PgD	-	=	OUT
'	1	2	3	4	5	6	7	8	9	0	DEL
TAB	q	w	e	r	t	y	u	i	o	p	ENT
SHF	a	s	d	f	g	h	j	k	l	;	SHF
	z	x	c	v	b	n	m	,	.	/	Mou
SL	CTL	ALT		SPACE					PSE	<-BACK	

Figura 7. Teclado, *The Eyegaze edge*. [6]

O *eye tracking* é feito pelo método da pupila/reflexão da córnea através de um *LED* infravermelho de baixa potência que ilumina o olho. Permite movimentos da cabeça num espaço de 3.8 x 3.0 x 3.8 cm. A câmara usada é sensível aos infravermelhos e capta 60 imagens por segundo. A calibração do sistema demora 15 segundos.

Apresenta um custo de aproximadamente 10.000 € sem extras. Os extras disponíveis são programas adicionais que permitem melhoria da síntese da fala ou o accionamento de luzes. [6,7]

1.2.2. **TM4, EYETECH DIGITAL SYSTEMS**

O *TM4* da *EyeTech Digital Systems* é direccionado para a tecnologia de assistência, amplamente usado por pessoas com atrofia muscular e esclerose lateral amiotrófica. Esta solução consiste num módulo de *eye tracking* com interface USB. É o sucessor do *TM3* que apresentava interface FireWire, o que limitava o número de computadores a que este módulo se podia conectar. Este módulo, apresentado na Figura 8, é usado para comunicar, enviar emails, navegar na internet, controlar o ambiente envolvente e jogar jogos.



Figura 8. *TM4, EyeTech Digital Systems.* [8]

Este sistema requer um computador com sistema operativo Microsoft Windows XP, Vista, ou 7 de 32-bit ou 64-bit, uma porta USB e um ecrã de 12 a 21 polegadas.

Para o *eye tracking* é utilizada uma câmara de infravermelhos, uma frequência de amostragem de 30 Hz e o método *dark pupil*. Permite movimentos da cabeça num espaço de 25 x 16 x 19 cm.

O custo desta solução ronda os 5.000 €, podendo subir até aos 10.000 € ao adicionar o suporte e o computador. [8,7]

1.2.3. *MYTOBII P10, TOBII*

Esta é uma das soluções mais populares entre as existentes. Criada pela *Tobii*, é direccionada para pessoas com incapacidades físicas. Como visível na Figura 9, este sistema é composto por um computador e um sistema de visão para o *eye tracking*.



Figura 9. *MyTobii P10, Tobii.* [9]

Neste sistema estão incluídos vários programas de utilização, tais como o envio de emails, jogos e editor de texto, podendo ser incluídas novas aplicações já existentes no computador.

A calibração do *MyTobii P10* demora aproximadamente 10 segundos. Para o *eye tracking* é utilizada uma câmara de infravermelhos com taxa de amostragem de 40 Hz e são utilizados dois métodos, o *dark pupil* e *bright pupil*. Permite movimentos da cabeça num espaço de 30 x 15 x 20 cm.

Encontram-se vários acessórios para este sistema, como por exemplo um suporte para secretária ou para cadeira de rodas e camas, módulos extra para o seguimento da cabeça (*head tracking*) ou para o controlo do meio envolvente através de comandos de infravermelhos. O preço vai desde os 12.000 € até aos 20.000 €. [9,7]

1.2.4. **EYEMAX, DYNAVOX**

Este sistema de tecnologia de assistência é desenhado para pessoas com incapacidades motoras. O *EyeMax* é um módulo que se liga a outro produto da *DynaVox*, o *Vmax*. Resultando no conjunto demonstrado na Figura 10.



Figura 10. *EyeMax* e *Vmax*, *DynaVox*. [10]

Pode ser usado por pessoas de qualquer idade e com capacidades diferentes, basta pelo menos conseguir usar um dos olhos. Esta solução tem várias aplicações, como demonstradas na Figura 11. Entre as habituais aplicações existentes nestes sistemas, destaca-se a aplicação do telefone, para atender ou fazer chamadas.



Figura 11. Aplicações, Vmax. [10]

Para o *eye tracking* é utilizada uma câmara de infravermelhos e o método *dark pupil*. Permite movimentos da cabeça num espaço de 15 x 23 x 29 cm.

Esta solução apresenta alguns sistemas de suporte, sendo facilmente integrada numa cadeira de rodas. O custo desta solução dependerá dos seus acessórios mas rondará os 12.000 €.[10,7]

1.2.5. INTELLIGAZE™, ALEA TECHNOLOGIES

Tipicamente utilizada por pessoas com paralisia cerebral, esclerose múltipla e esclerose lateral amiotrófica, esta é uma solução que permite controlar aplicações do *Windows* através de *eye tracking*. O *Intelligaze™* é constituído por um módulo de visão, com interface USB, apresentado na Figura 12.

Entre as aplicações já existentes no computador, são instaladas também algumas aplicações tais como teclado, jogos e síntese de fala. Este sistema requer um computador com sistema operativo *Windows XP* ou *Vista* com monitor entre 12 e 19 polegadas. Este módulo traz consigo um sistema de encaixe para o monitor.

Para o *eye tracking* é utilizada uma câmara infravermelhos com frequência de amostragem de 50 Hz e os dois métodos (*bright* e *dark pupil*). Permite movimentos da cabeça num espaço de 30 x 20 x 20 cm.



Figura 12. *Intelligaze™, alea technologies.*[11]

O custo desta solução dependerá das actualizações de software e sistemas de integração (suportes de montagem), rondando os 6.000 €. [11,7]

1.2.6. SEETECH®, HUMANELEKTRONIK GMBH

O *SeeTech®* foi especialmente desenvolvido para pessoas com deficiência motora. O sistema é composto por um computador, um módulo de visão e um suporte, como visível na Figura 13.



Figura 13. *SeeTech®, Humanelektronik GmbH.* [12]

Permite interagir com quase todas as aplicações *Windows* e ainda algumas disponibilizadas pelo próprio sistema, tais como teclado, jogos, síntese de fala e controlo de meio envolvente.

Para o *eye tracking* é usada uma câmara infravermelhos e o método *dark pupil* a uma frequência de 40 Hz. Permite movimento da cabeça num espaço de 30 x 20 x 20 cm.

Dependendo da configuração, esta solução apresenta preços entre os 7.000 e os 10.000 €. As configurações possíveis incluem módulos Bluetooth, joystick, bateria recarregável para a cadeira de rodas, entre outros. [12,7]

1.2.7. ECO2 com ECOPOINT, PRENTKE ROMICH COMPANY

Constituído por um computador (*ECO2*) e um módulo de visão (*ECOpoint*), visível na Figura 14, este é um sistema desenhado especialmente para pessoas com autismo mas também pode ser utilizado por pessoas com incapacidades motoras. Apesar de o módulo *ECOpoint* ser vendido em separado, este só pode ser usado com outros produtos da *Prentke Romich Company*.



Figura 14. *ECO2 com ECOpoint, Prentke Romich Company*. [13]

As aplicações de interação existentes são as habituais contudo, graficamente, são bastante diferentes, pois a interface é baseada em figuras em vez de palavras. Só está disponível em inglês.

Para o *eye tracking*, é usada uma captura de imagem com frequência de 40 Hz e permite movimentos da cabeça num espaço de 25 x 38 x 20 cm.

Esta é uma solução que dispõe de várias configurações de suporte para cadeiras de rodas, para chão, para cama, etc., cada uma tem um custo, indo dos 11.000 € até aos 15.000 €. [13,7]

1.2.8. *EYEWRITER*

O *EyeWriter* é um projecto, ainda em andamento, de pesquisa, aplicado a pessoas com esclerose lateral amiotrófica. A equipa que compõe o *EyeWriter* trabalha em conjunto com membros do *Free Art and Technology* (FAT), *OpenFrameworks* e *Graffiti Resarch Lab*, para criar um sistema de *eye tracking* de baixo custo que possibilita as pessoas desenharem com os olhos. Na Figura 15 é apresentada esta solução.



Figura 15. *EyeWriter*. [14]

Apesar de o *EyeWriter* não estar de acordo com os objectivos deste trabalho em relação à sua configuração, é uma solução que mostra muito interesse devido ao seu livre acesso, isto é, no geral qualquer pessoa pode recriar o *EyeWriter* apenas precisando de ter o hardware e software necessários e seguir as instruções de montagem, tudo disponível no site da organização. [14]

Para o *eye tracking* é utilizada a configuração apresentada na Figura 16 e é usado o método *dark pupil*.

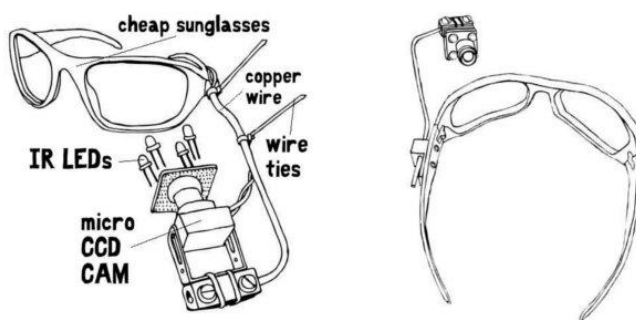


Figura 16. Configuração do *EyeWriter*. [14]

A solução apresenta um custo de aproximadamente 50\$ (USD) tendo como pré-requisito um computador. [14]

1.3. TABELA COMPARATIVA

De seguida é apresentada a Tabela 1 que reúne, de maneira resumida, a informação encontrada no estado da arte. Esta tabela também facilita a comparação entre os diversos produtos.

Tabela 1. Tabela resumo.

Produto	Preço base	Aplicações base	Área de movimento (cm)	Requisitos
The Eyegaze Edge™	10.000 €	Teclado, Jogos, Frases, Síntese de fala.	3,8 x 3 x 3,8	-
TM4	5.000 €	Teclado, Jogos, Internet, Controlo ambiente, Síntese de fala.	25 x 16 x 19	Computador com Windows XP ou Vista
MyTobii P10	12.000 €	Teclado, Jogos, Email	30 x 15 x 20	-
EyeMax	12.000 €	Teclado, Jogos, Síntese de fala, Telefone	15 x 23 x 29	-
IntelliGaze™	6.000 €	Aplicações Windows, Teclado, Jogos, Síntese de fala	30 x 20 x 20	-
SeeTech	7.000 €	Aplicações Windows, Teclado, Jogos, Síntese de fala, Controlo ambiente	30 x 20 x 20	-
ECO2 com ECOpoint	11.000 €	Teclado, Jogos, Síntese de fala, Controlo ambiente	25 x 38 x 20	-
EyeWriter	50 \$(USD)	Ferramenta de desenho	-	Computador

2. SELECÇÃO DE HARDWARE E SOFTWARE

Da análise do capítulo anterior é possível indicar quais as características mais importantes, tanto de hardware como de software, a ter em conta no desenvolvimento de um aparelho de *eye tracking* utilizado por pessoas com tetraplegia. Estas características são:

- Câmara de infravermelhos com alto débito de frames por segundo;
- Luz infravermelha;
- Computador e ecrã;
- Estrutura de apoio para cama ou cadeira de rodas;
- Interface simples e directa.

Em relação ao hardware, uma vez que neste caso o utilizador não mexe a cabeça não é necessária uma câmara com resolução muito alta. Isto porque o olho vai ser filmado directamente sem a necessidade de primeiro localizar a face e consequentemente o olho. As propriedades mais importantes da câmara são a sua capacidade de captar infravermelhos e o seu débito de frames por segundo, ou *frame rate*, em inglês.

A câmara escolhida foi a *PlayStation® Eye* (Figura 17) devido ao seu custo reduzido, alto *frame rate* (60 fps a uma resolução de 640×480 pixéis) e compatibilidade com o sistema operativo utilizado. Para tornar esta câmara sensível aos infravermelhos (IV) foi trocado o filtro IV existente na câmara por um filtro de onda visível. Isto é suficiente pois o sensor de vídeo é sensível aos raios IV.



Figura 17. *PlayStation® Eye*. [15]

Um dos outros motivos que levou à escolha desta câmara foi o facto de existir muita bibliografia sobre a conversão da *PlayStation® Eye* numa câmara de infravermelhos [16]. Foi também trocada a lente, uma vez que a de origem tinha uma distância focal muito pequena (6 mm), sendo colocada uma lente de 18 mm. A diferença é mostrada na Figura 18, em que à mesma distância da câmara, são obtidas imagens diferentes.



Figura 18. Direita: 18 mm; esquerda: 6 mm.

Como fonte de luz IV foram escolhidos *leds* infravermelhos na gama *NIR* (*near-infrared*, do português próximo de infravermelho), aproximadamente 750 nm, pois quanto menor o comprimento de onda, menores são os danos para o olho. Os *leds* foram adquiridos já soldados a uma placa. O conjunto tem o nome de iluminador IV e está visível na Figura 19.



Figura 19. Iluminador infravermelho de 12 led.

Quanto ao computador e ecrã, para o desenvolvimento do projecto foi utilizado um portátil *Toshiba Satellite L650*, mas o sistema pode facilmente ser adaptado a outros computadores. O ecrã deve ter, no mínimo, 15,6''. Porque ao utilizar um ecrã menor, a dimensão dos botões diminui, assim

como o movimento dos olhos é menor, diminuindo assim a resolução, podendo o sistema ficar inutilizável.

Uma preocupação constante é o custo do equipamento no seu todo, logo, tanto o sistema operativo como as ferramentas de software necessárias são *open source*. Assim, o sistema operativo escolhido foi o Linux na distribuição *Ubuntu* versão 10.04, o ambiente de desenvolvimento foi o *Qt Creator IDE* versão 4.7 e foi também utilizada uma biblioteca para a visão computacional em tempo real, o *OpenCV* (*Open Source Computer Vision Library*), versão 2.2 [17].

3. EYE TRACKING

Para o *eye tracking* foi utilizado o método *dark pupil* não só porque é o mais utilizado entre os exemplos vistos anteriormente mas, principalmente, porque o método *bright pupil* é mais eficaz em meios escuros ou com uma fonte de luz IV mais potente. Como é visível na Figura 20, para as mesmas condições de luz, o método *bright pupil* não produz uma pupila clara o suficiente para que seja facilmente detectada, enquanto o *dark pupil* produz uma pupila bastante escura. Também, uma vez que o utilizador tem restrições motoras, o sistema será utilizado primordialmente em interior (dentro de casa/edifício) com luz controlada, tornando mais vantajoso o método *dark pupil*.

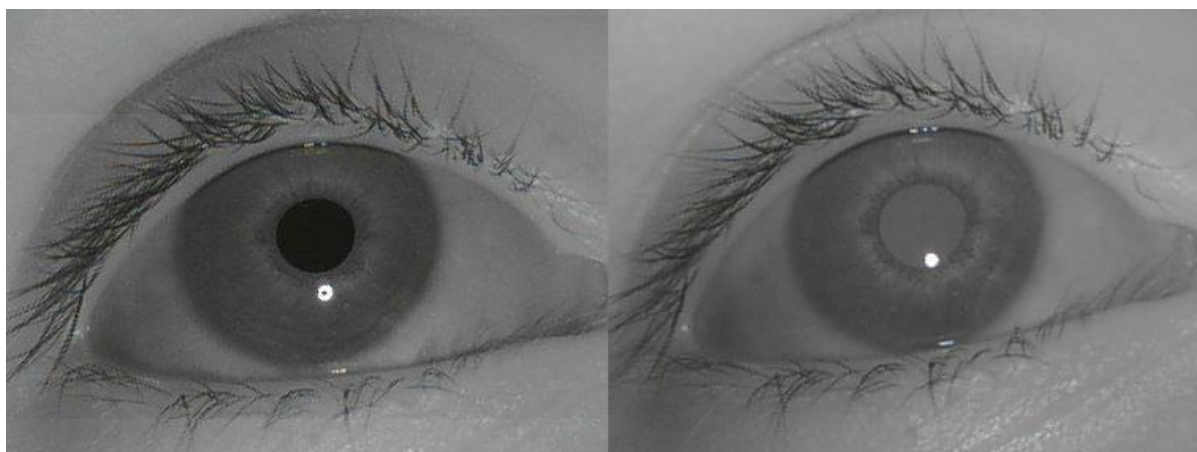


Figura 20. *Dark pupil* (esquerda) e *bright pupil* (direita).

O primeiro passo para o *eye tracking* é detectar a pupila. Depois, os dados daí provenientes são interpretados e faz-se corresponder o olhar com um ponto no ecrã. De seguida, são aprofundadas todas as técnicas usadas para o *eye tracking*, e posteriormente é apresentado o fluxograma.

3.1. DETECÇÃO DA PUPILA E DO GLINT

A detecção de objectos numa imagem pode ser feita através da cor/luminosidade, através da forma do objecto ou ainda através de uma junção das duas propriedades. Qualquer um dos métodos necessita de processamento prévio da imagem como, por exemplo, redução de ruído. De seguida, os dois métodos são apresentados.

3.1.1. COR/LUMINOSIDADE

Apesar de a imagem infravermelha obtida ser uma imagem em tons de cinza, pode-se falar da luminosidade de cada pixel. Diz-se que um pixel é brilhante quanto mais claro ele for. A pupila, neste caso, é escura e este pode ser um dos meios para a sua detecção.

Recorre-se á técnica de *blob detection* (detecção de regiões) que consiste na etiquetagem de regiões com as mesmas propriedades de cor/luminosidade. O algoritmo de detecção de *blobs* está descrito no Apêndice I. De uma maneira resumida, é percorrida toda a imagem e são colocadas *labels* (etiquetas) em todos os pixéis dependendo da sua luminosidade e proximidade. A um conjunto de *labels* iguais dá-se o nome de *blob*.

Para muitos casos isto será suficiente mas neste não, porque, mesmo retirando ruído da imagem através do *Gaussian blur*, as pestanas, cantos e, por vezes, as sobrancelhas produzem *blobs*, que são interpretados como falsos positivos da pupila. Uma forma de determinar se um *blob* corresponde à pupila é verificar se a área desse *blob* é aproximada á área da pupila.

Existem, contudo, vários contratempos nesta técnica quando aplicada a este projecto. Uma imagem, ao conter centenas de *blobs*, eleva demasiado o tempo de processamento para a comparação de áreas. Uma comparação de área mostrou também não ser suficiente pela razão de existirem vários *blobs* com dimensão aproximada e a área da pupila não ser constante.

3.1.2. FORMA

Quanto à forma, existem várias técnicas de detecção, sendo umas orientadas à detecção de rectas, outras de elipses, outras de formas angulosas, etc. Para a detecção de círculos (entre outros) existe a transformada de Hough. A transformada de Hough baseia-se no agrupamento de pontos do objecto através de um procedimento de votação. Consiste, portanto, na transformação de uma imagem num acumulador de votos.

Antes da votação é preparada a imagem. São detectados os contornos de maneira a que a transformada seja aplicada correctamente. Os contornos são detectadas a partir de mudanças bruscas de luminosidade ou por descontinuidade. Para tal, é usado o operador *canny* que faz parte da biblioteca *OpenCV*. A imagem obtida após o operador *canny* é apresentada na Figura 21. É uma imagem de um canal com 8 bits, ou seja, em tons de cinza, mas com a particularidade de só conter pixéis brancos ou pretos.

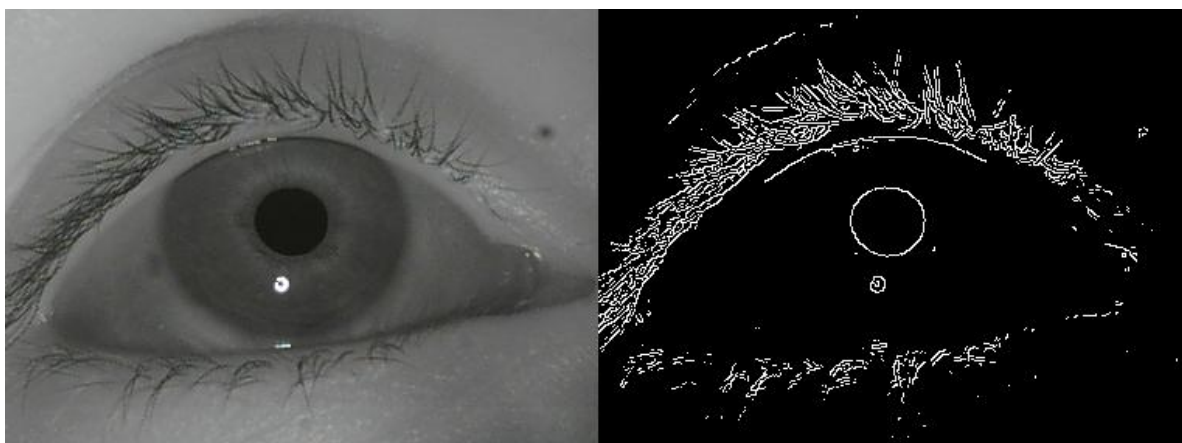


Figura 21. Resultado do operador *canny*.

É sobre esta imagem que as votações são realizadas. Percorre-se a imagem, e quando é encontrado um pixel branco, é realizada a votação numa matriz paralela à imagem (do mesmo tamanho que a imagem). Como a forma a detectar é um círculo, os votos são realizados à volta das coordenadas do pixel branco a uma distância r (raio). Na Figura 22 estão apresentadas duas imagens, do lado esquerdo tem-se a imagem *canny* e do lado direito a matriz de votos.

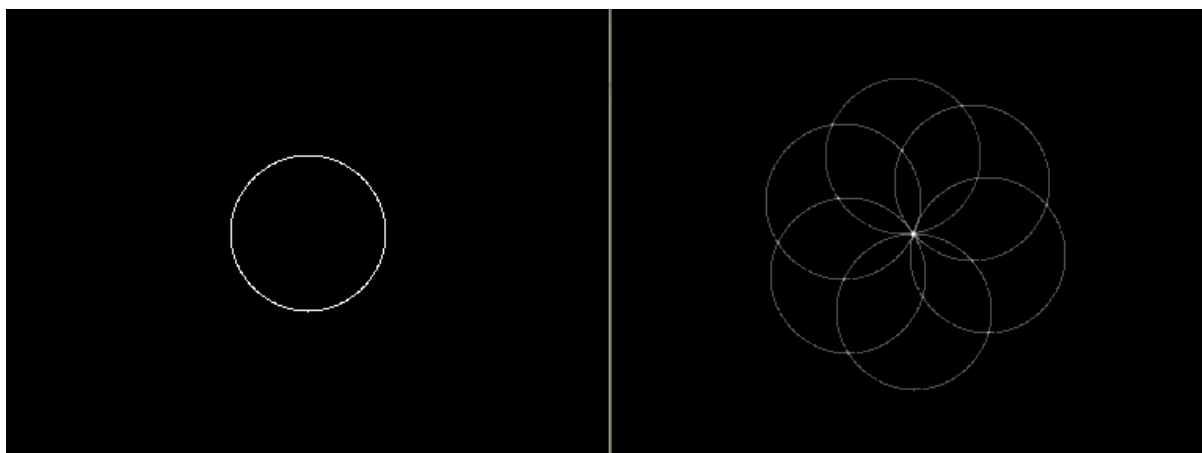


Figura 22. Transformada de *Hough*.

Por motivos de tempo de processamento os votos são feitos para ângulos com intervalos de 10° , sem prejudicar o resultado. Pode verificar-se que, depois de realizados os votos, existem elementos da matriz com maior número votos, como visível na Figura 23 através do pico, com 12 votos.

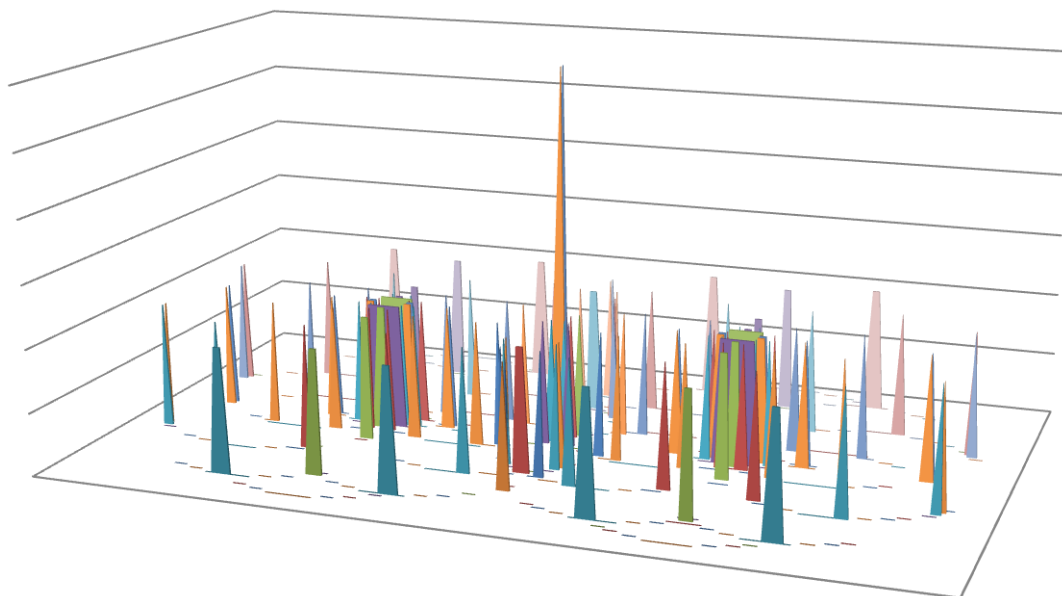


Figura 23. Gráfico da matriz de votos.

É esperado obter um máximo local mas, na maioria das vezes, são encontrados pelo menos dois máximos. Isto deve-se ao facto de o raio utilizado na votação não ser igual ao raio do círculo que se pretende detectar.

É importante conhecer o raio do círculo que se procura, pois podem-se obter resultados como o mostrado na Figura 24. À esquerda foi usado um raio superior ao círculo a detectar e à direita foi usado um raio inferior.

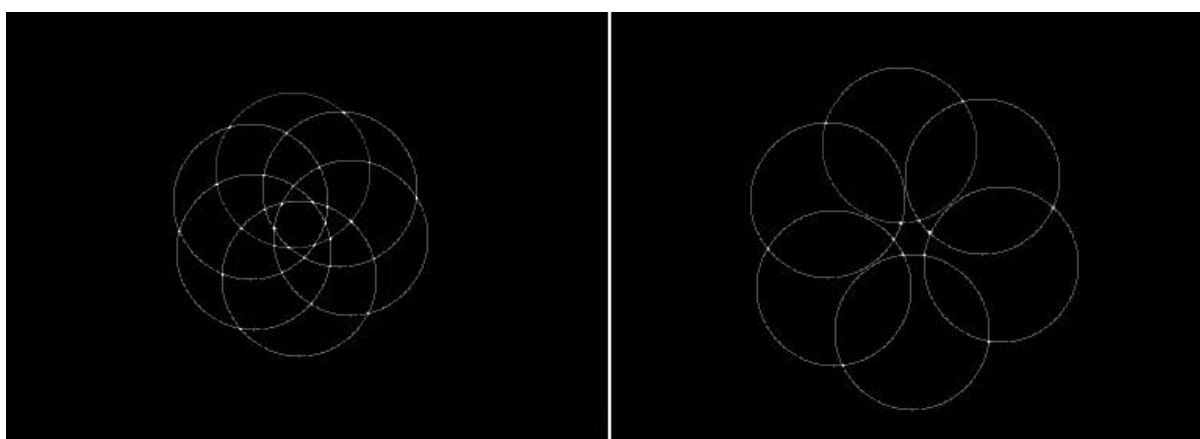


Figura 24. Transformada de Hough com raios diferentes.

Neste projecto, o raio da pupila pode apresentar-se como um problema. A pupila raramente se encontra com um raio estável, dependendo da iluminação ambiente e da luminosidade do objecto observado. Por exemplo, num ambiente com luz constante e controlada, podem ser registados raios

de pupila como os apresentados na Figura 25, consoante as características de luminosidade do ponto observado.

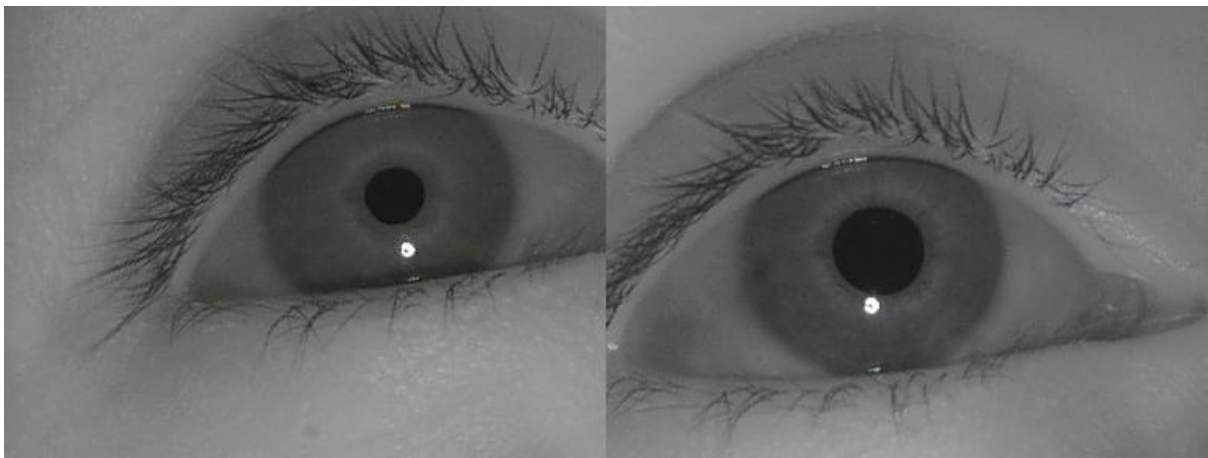


Figura 25. Esquerda: olho quando se olha para algo claro. Direita: olho quando se olha para algo escuro.

A solução deste problema reside na realização de várias matrizes de voto que correspondem a vários raios de pupila. As votações são feitas para 12 raios desde os 18 pixéis até aos 30 pixéis. Desta forma é possível detectar sempre o centro da pupila.

Pelo mesmo método é detectado o *glint*, como visível na Figura 26. Mas desta vez não são necessárias várias matrizes de voto dado o raio do *glint* ser constante, já que o utilizador não move a cabeça no sentido de se aproximar ou afastar da câmara. Contudo, a transformada de *Hough* para raios muito pequenos, como neste caso, pode apresentar alguns erros, sendo estes corrigidos ao aumentar o intervalo entre os ângulos de voto para 20°.

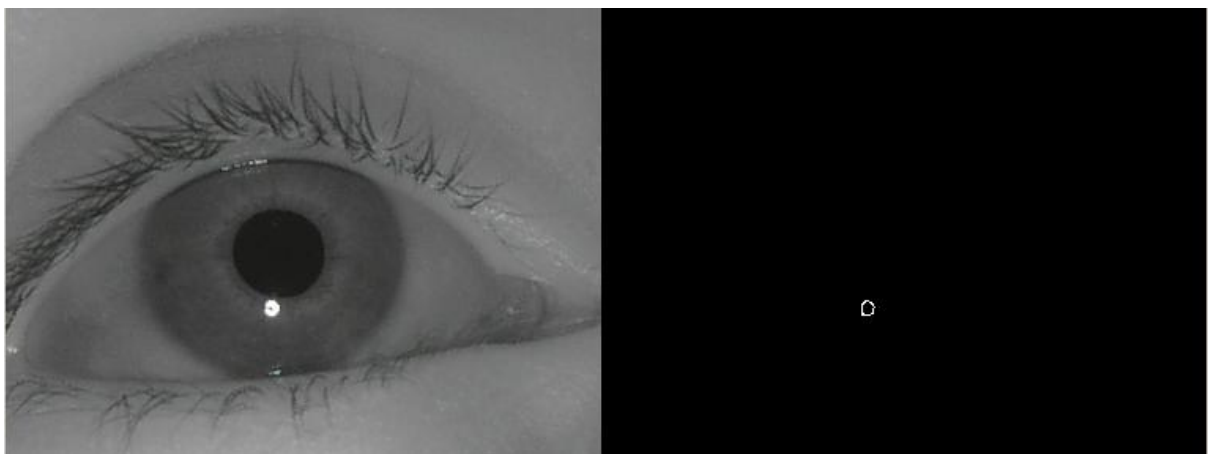


Figura 26. Detecção do *glint*.

3.2. RECOLHA E TRATAMENTO DE DADOS

Os dados recolhidos para a interpretação do problema e obtenção de resultados sofreu mudanças ao longo do projecto. Numa primeira etapa foram só recolhidas as coordenadas do centro da pupila, pois considerou-se que, pelo facto de o utilizador não ser capaz de mover a cabeça, o centro da pupila podia ser relacionado com um ponto de referência estático (o ponto (0,0) por exemplo). Nos primeiros testes observou-se que mesmo os movimentos respiratórios do utilizador alteravam o ponto de referência, inutilizando a calibração efectuada e impossibilitando o normal funcionamento do sistema.

A solução foi ter um ponto de referência móvel, o *glint*. Agora, o dado recolhido é o centro da pupila em relação ao *glint*, como visível na Figura 27, que apresenta valores não influenciados pelos pequenos movimentos da cabeça.

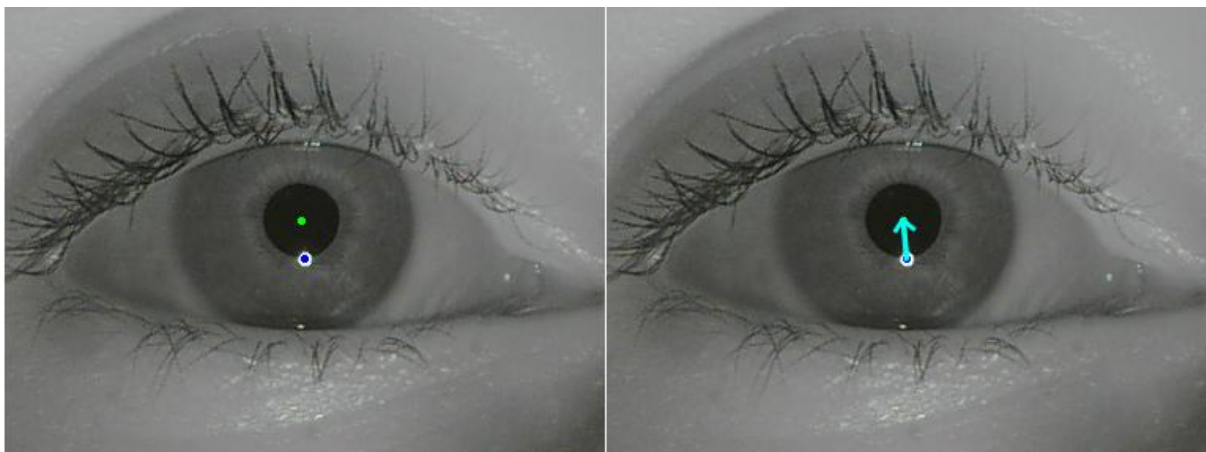


Figura 27. Vector entre o *glint* e a pupila.

Como o utilizador tem incapacidades motoras, presume-se que a distância entre o utilizador e a câmara mantém-se constante. Se isto não fosse verdade, este seria outro parâmetro a ter conta.

3.2.1. FILTRO DE *KALMAN*

Os dados obtidos não podem ser usados directamente, uma vez que a constante alteração do raio da pupila, que apesar de ser pequena agrava-se por causa da baixa resolução de imagem, pode alterar o centro da pupila.

O filtro de Kalman é um estimador óptimo que prevê e corrige o estado estimado de um sistema. Consiste na aplicação de dois grupos de equações. O primeiro grupo, chamado de *time*

update (actualização temporal), é responsável pelo avanço no tempo da previsão, onde o erro da covariância estima *a priori* o próximo estado. Constituído por duas equações, na equação (1) estima-se o estado onde \hat{x}_k^- é o estado estimado, A é a matriz do sistema, \hat{x}_{k-1} é o estado estimado anterior, B é a matriz de entrada e u_k é o vector de controlo. Na equação (2) é estimado o erro associado à estimação de estado, onde P_k^- é a covariância, obtida a partir da covariância calculada no *measurement update* (P_{k-1}), e a covariância do ruído do processo, Q .

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k \quad (1)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (2)$$

O segundo grupo, chamado de *measurement update* (actualização de medida), incorpora uma nova medição no estado estimado *a priori* para corrigir a estimação *a posteriori*. Constituído por três equações, a equação (3) calcula o ganho de kalman onde R é a matriz de covariância do ruído de medida. Depois, a equação (4) calcula o estado corrigido através da nova medição z_k . E, finalmente, a equação (5) actualiza o valor da covariância.

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (3)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (4)$$

$$P_k = (I - K_k H)P_k^- \quad (5)$$

A afinação do filtro é feita através das variáveis de ruído de processo e de medida. Nas Figura 28, Figura 29, Figura 30 e Figura 31 mostram-se as diferenças no comportamento do filtro em relação ao valor das variáveis. A resposta mais conveniente é a que reage mais rapidamente mas que seja pouco sensível ao ruído, então, os valores escolhidos para as variáveis foram os apresentados na Figura 30.

A biblioteca *OpenCV* é também usada na implementação deste filtro.

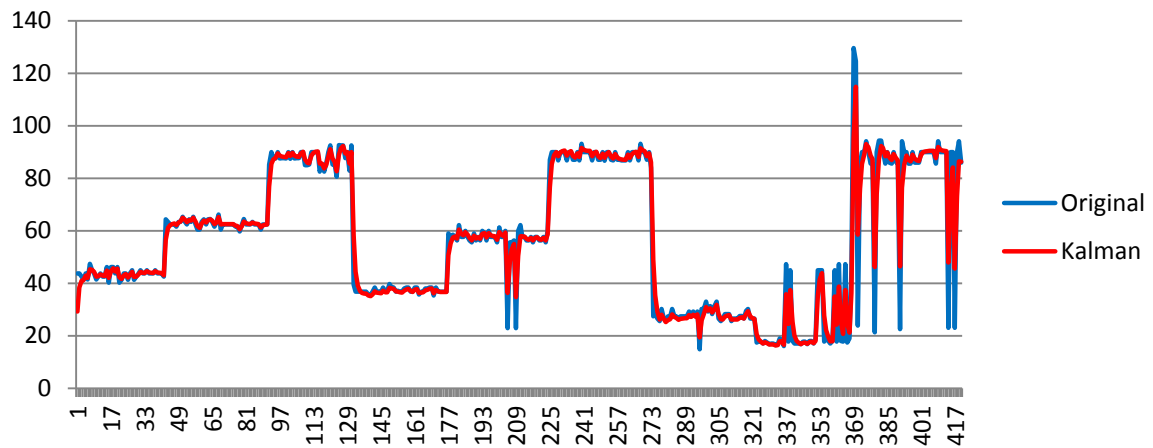


Figura 28. Resposta do filtro de kalman, com $Q = 1 \times 10^{-1}$ e $R = 0,1$.



Figura 29. Resposta do filtro de kalman, com $Q = 1 \times 10^{-2}$ e $R = 0,1$.

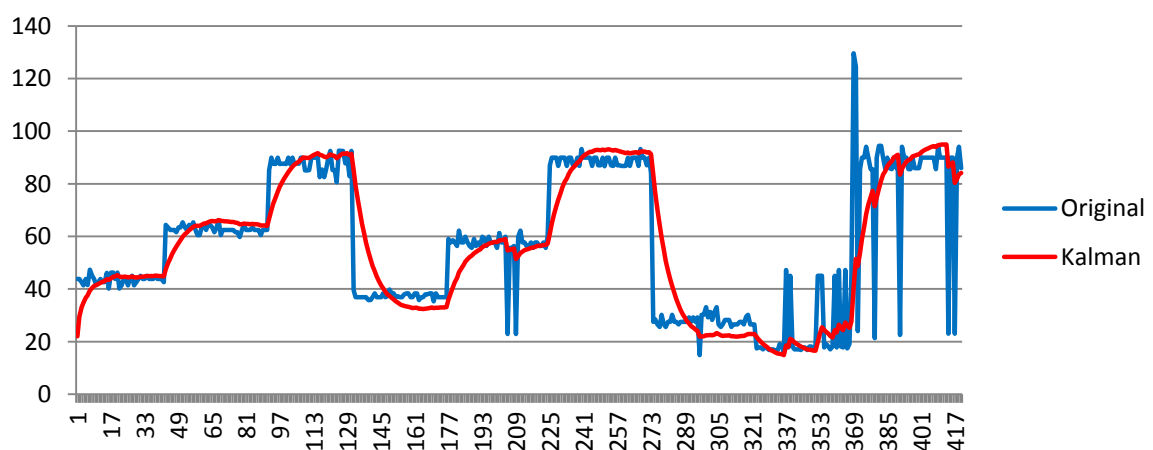


Figura 30. Resposta do filtro de kalman, com $Q = 1 \times 10^{-3}$ e $R = 0,1$.

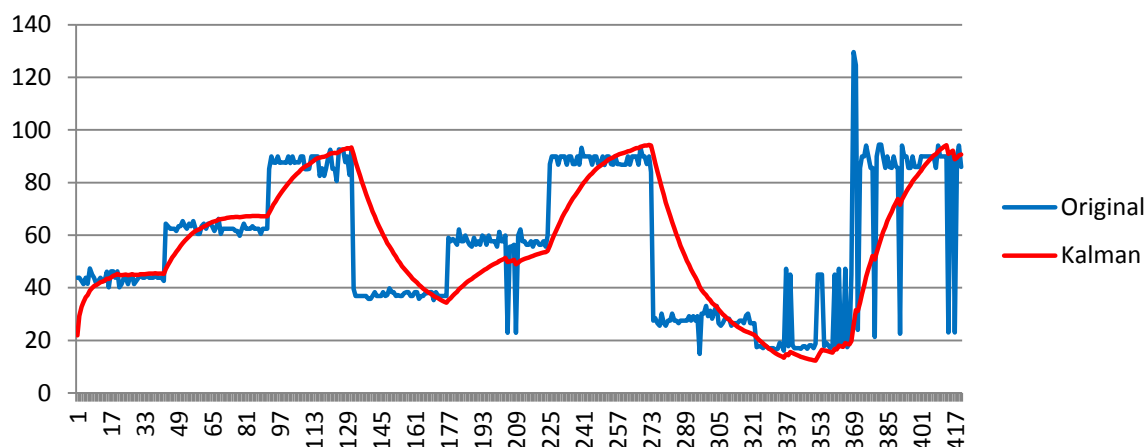


Figura 31. Resposta do filtro de kalman, com $Q = 1 \times 10^{-4}$ e $R = 0,1$.

3.3. MAPEAMENTO

Depois de se detectar a pupila e tratar os dados, é necessário fazer corresponder os pontos obtidos da pupila aos pontos do ecrã. A isto chama-se mapeamento. Para isto são necessários vários pontos referência, para que depois sejam a base do cálculo de todas as correspondências possíveis. À primeira parte dá-se o nome de calibração, que recolhe os dados. Na segunda parte são feitos os cálculos necessários para que se faça a correspondência entre o olho e os pontos do ecrã.

3.3.1. CALIBRAÇÃO

A calibração consiste na recolha dos dados da pupila que se fazem corresponder aos pontos fixos do ecrã, sendo estes os pontos de referência ou de calibração. Numa primeira abordagem foram utilizados 9 pontos de calibração. Contudo, os dados provenientes da pupila não formam rectângulos, como o conjunto de coordenadas do ecrã, como é visível na Figura 32.

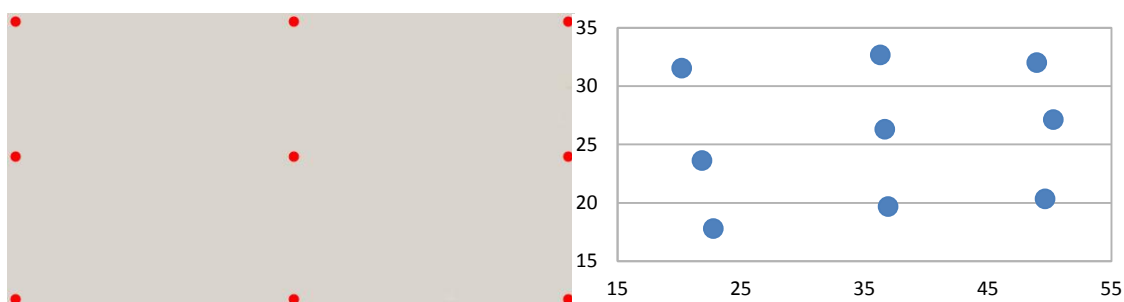


Figura 32. Pontos do ecrã (em cima). Pontos da pupila (em baixo).

Existem várias soluções para este problema. Uma delas é fazer uma regressão polinomial dos pontos na horizontal ou vertical. Outra solução, mais simples de calcular, é fazer uma regressão linear, mas aumentando o número de pontos de calibração para minimizar o erro. Optou-se por fazer uma regressão linear com 20 pontos de calibração, obtendo-se assim um possível resultado visível na Figura 33.

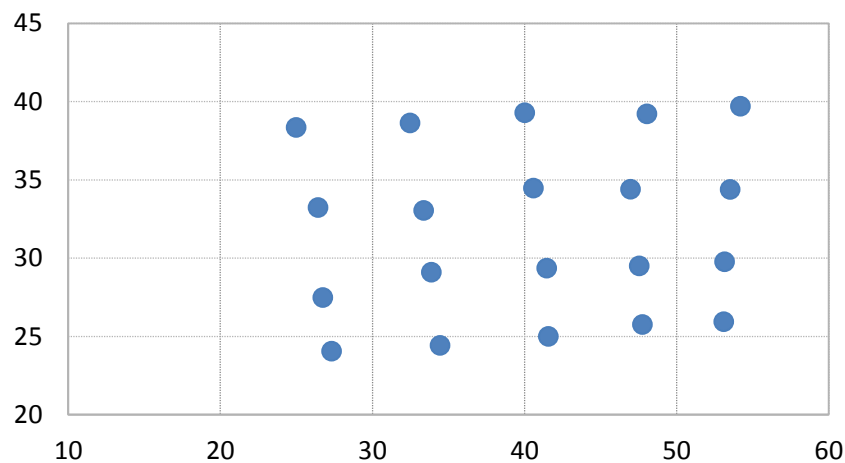


Figura 33. Coordenadas da pupila.

3.3.2. TRANSFORMAÇÃO DE COORDENADAS

Da calibração resultam 20 pontos da pupila que formam 12 quadrantes. Todos os pontos formam rectas com os pontos mais próximos, como descrito na Figura 34. Daqui são guardados os parâmetros a e b de todas as rectas, onde a é o declive da recta e b é o ponto em que a recta intercepta o eixo dos yy .

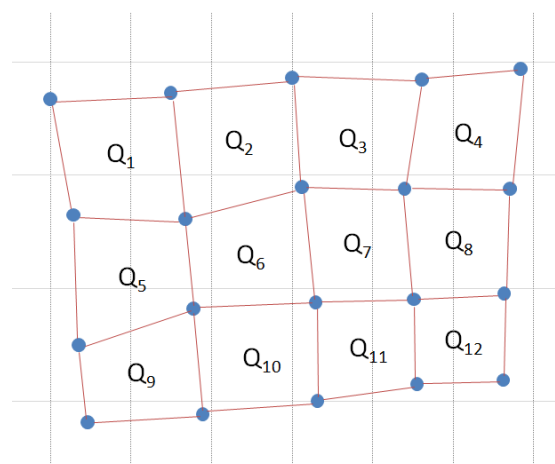


Figura 34. Rectas e quadrantes.

O primeiro cálculo é feito para perceber em que quadrante se encontra o novo ponto da pupila. Para isso verifica-se se o ponto está acima ou abaixo de cada uma das rectas horizontais e se está à esquerda ou à direita de cada uma das rectas verticais. De seguida, é feita uma relação entre o ponto da pupila e as rectas relativas ao quadrante, como visível na Figura 35. Essa relação é transposta para o quadrante do ecrã, obtendo-se quatro pontos que formam duas rectas, uma horizontal e outra vertical, como descrito na Figura 36. O ponto de intercepção das rectas é o ponto do cursor, também descrito na Figura 36 como ponto C.

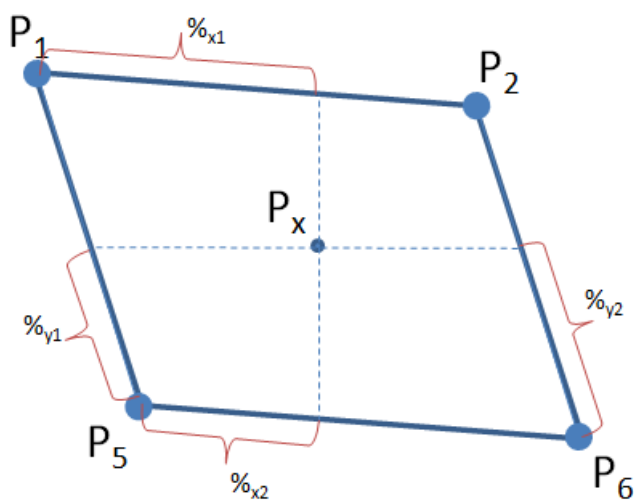


Figura 35. Relação da pupila.

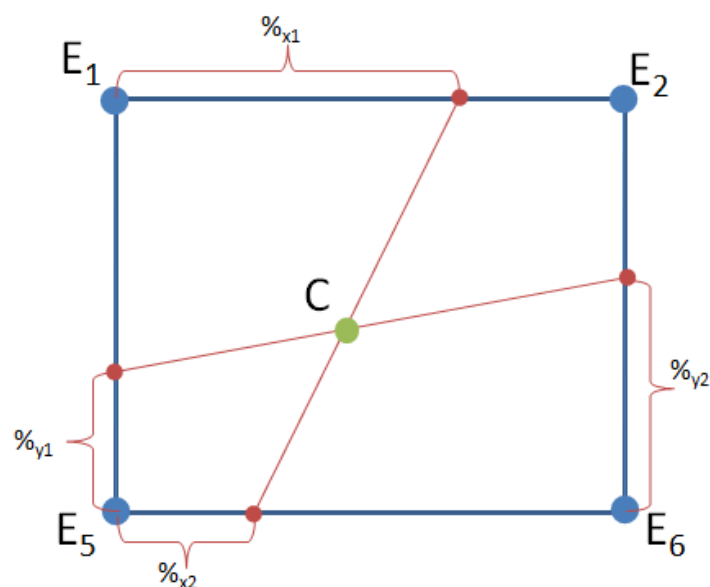


Figura 36. Relação do ecrã.

3.3.2.1. Resolução

A resolução do sistema desenvolvido até ao momento será importante para a criação da interface pois ditará o tamanho mínimo dos botões a apresentar. Existe a resolução em x e em y .

A resolução em x é o movimento horizontal do olho equivalente ao comprimento de ecrã, ou seja, tendo o ecrã 1366 pixéis na horizontal e o olho se mover 29 pixéis ao longo deste, o movimento de um pixel do olho corresponde a 47 pixéis do ecrã.

A resolução em y é o movimento vertical do olho equivalente à largura do ecrã. Analogamente à resolução em x , o ecrã tem 768 pixéis na vertical que correspondem a 14 pixéis de movimento do olho, ou seja, cada pixel movido na vertical pelo olho, equivalem a 54 pixéis do ecrã.

A resolução é influenciada por vários motivos, tais como a distância entre o utilizador e o sistema, que neste caso foram de 60 cm do ecrã e 30 cm da câmara, aproximadamente, e o tamanho do ecrã, sendo utilizado um de 15,6''. Contudo existe um factor que não se pode controlar, a calibração.

Para o cálculo da resolução foi usada a calibração mostrada anteriormente na Figura 33.

3.4. FLUXOGRAMA

Na Figura 37 é apresentado o fluxograma do *eye tracking*.

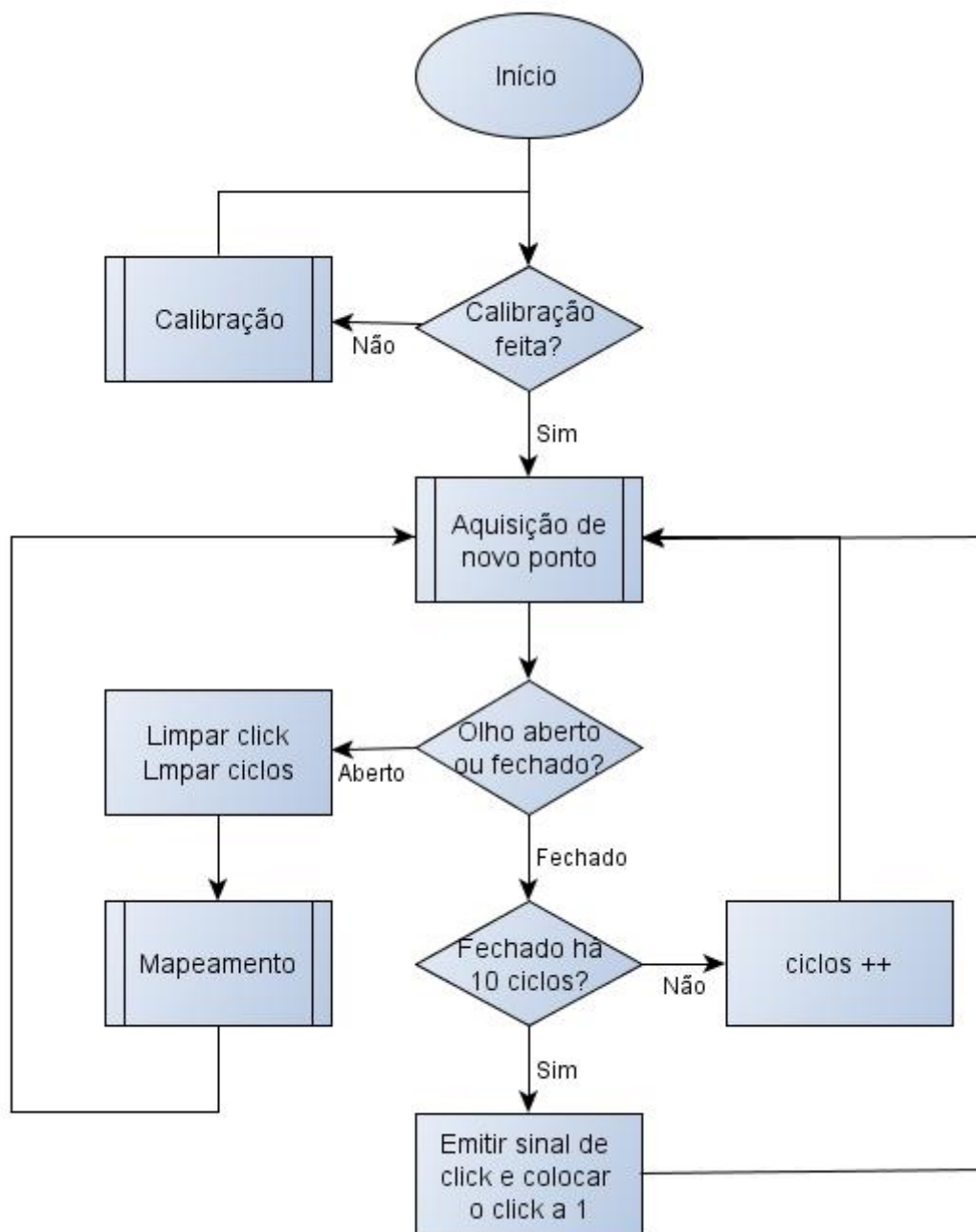


Figura 37. Fluxograma de *eye tracking*.

4. INTERFACE

A interface deve adequar-se ao utilizador. Neste caso, deve ser uma interface simples e útil, tendo como principal objectivo melhorar a capacidade de comunicação do utilizador. A existência de muitos factores físicos e ambientais variáveis, faz com que seja necessário efectuar uma calibração cada vez que se inicia o programa. As aplicações são apresentadas de seguida.

O menu principal, apresentado na Figura 38, mostra as 5 aplicações principais do programa. O último botão, o botão “Sair”, fecha o programa, mostrando um ecrã de alerta para confirmar o fecho. No Apêndice II é apresentado o manual de utilização do sistema.

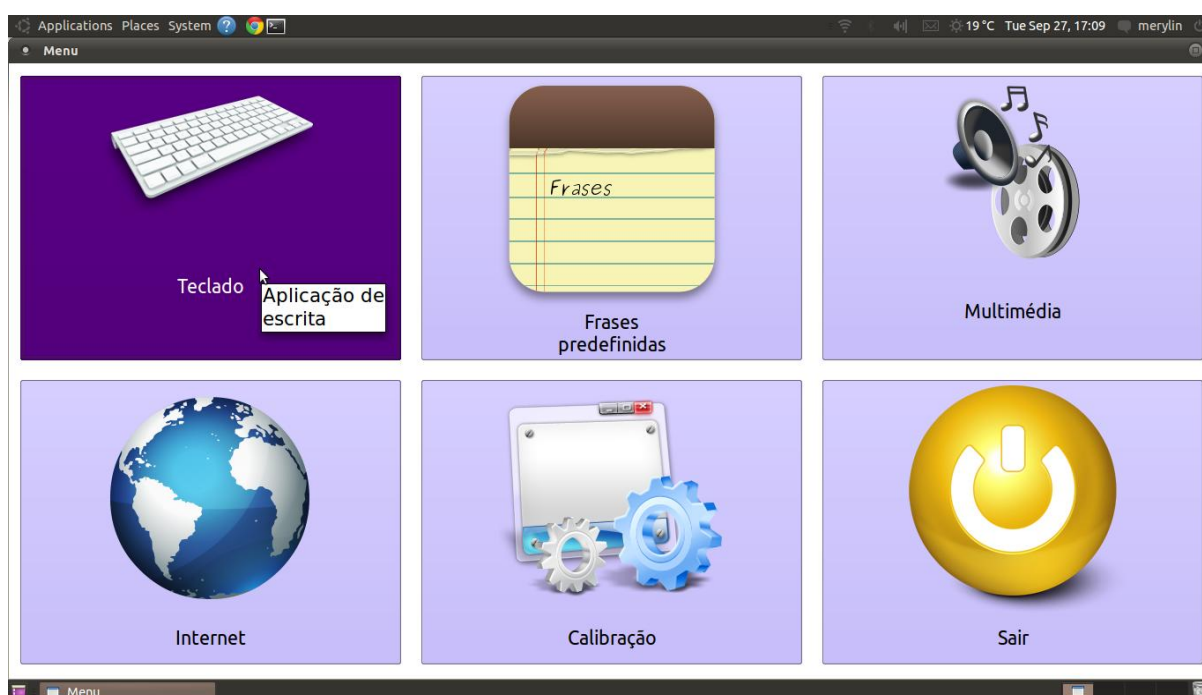


Figura 38. Menu.

4.1. TECLADO

O Teclado é uma das aplicações mais importantes do sistema. Tem todas as teclas de um teclado normal destinadas para escrita. Na Figura 39 apresenta-se o ambiente gráfico desta aplicação.

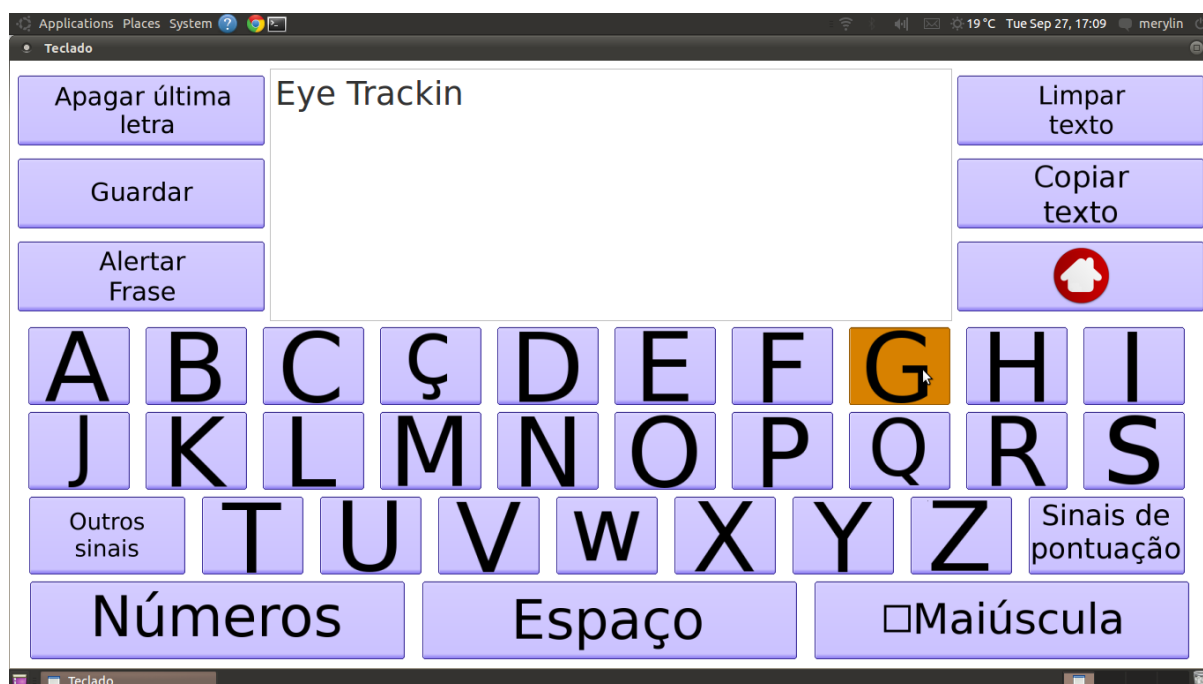


Figura 39. Teclado.

Nesta aplicação existem botões com finalidades particulares. O botão “Guardar” guarda o texto existente num ficheiro com extensão *.txt*. O botão “Alertar Frase” reproduz um som de alerta e apresenta o texto no ecrã como visível na Figura 40.

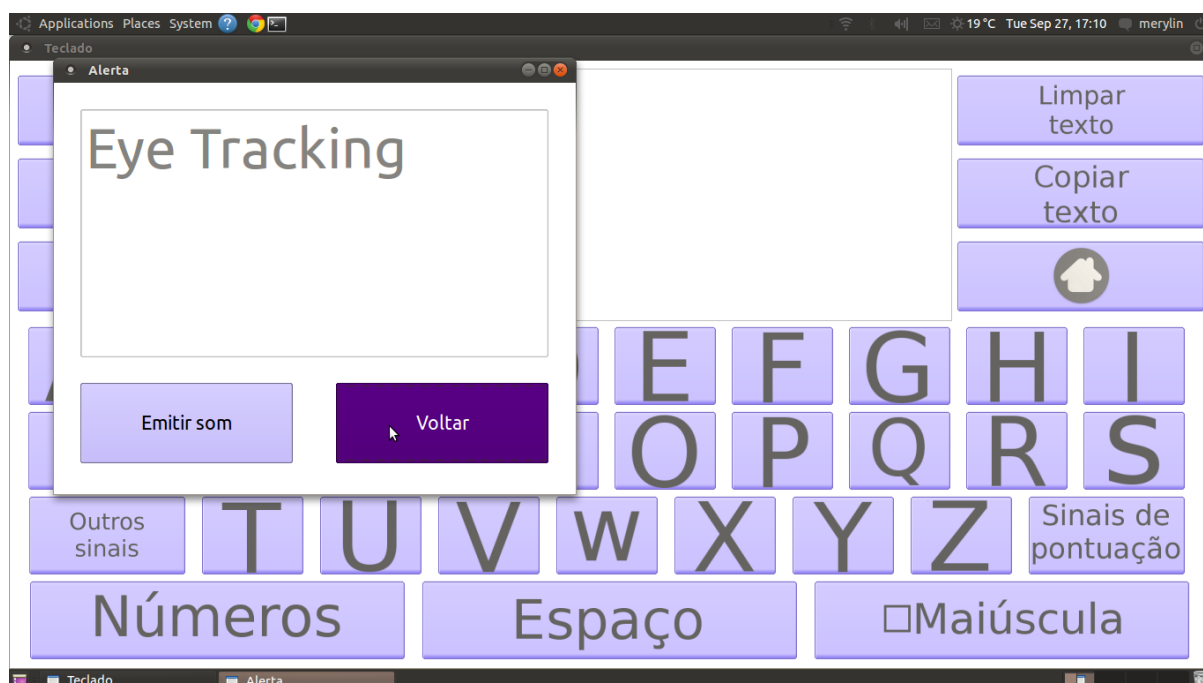


Figura 40. Alerta.

4.2. FRASES PREDEFINIDAS

Esta aplicação tem a finalidade de reduzir o tempo de escrita do utilizador. Contém 14 frases pré-definidas pelo utilizador de maneira a serem usadas de imediato. Estas frases podem ser editadas pelo utilizador a qualquer momento. Em baixo, na Figura 41, é apresentada a interface desta aplicação.

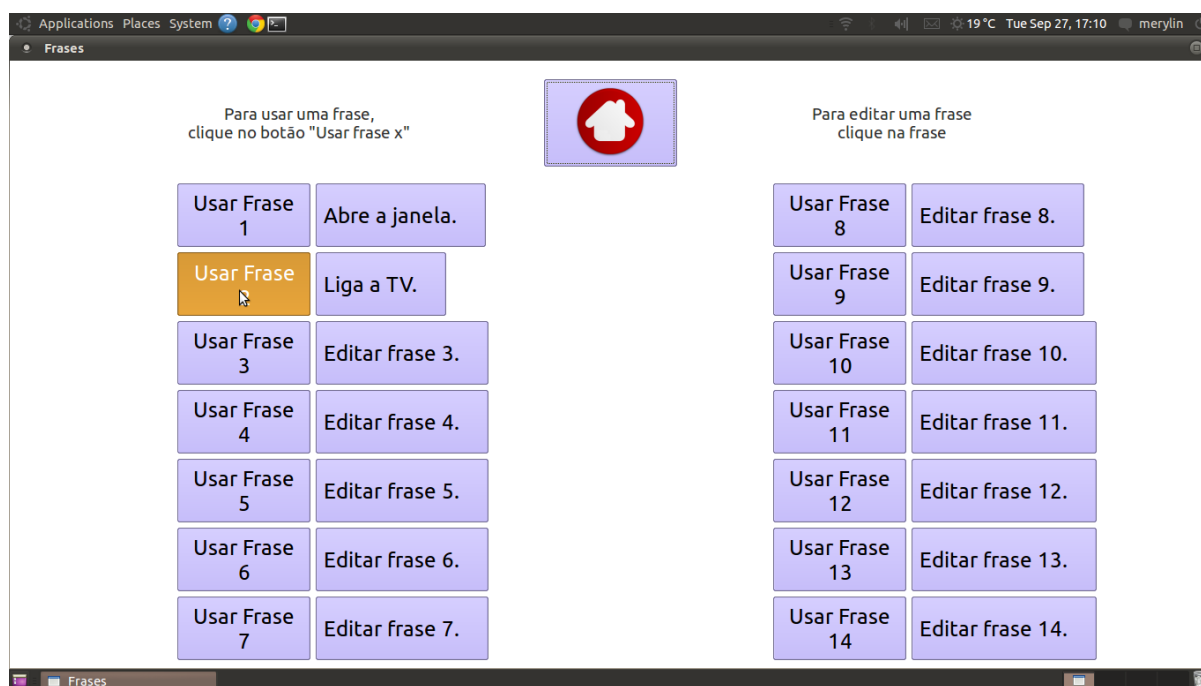


Figura 41. Frases predefinidas.

Nesta aplicação, as frases podem ser editadas através do clique na frase, ou podem ser usadas ao clicar no botão “Usar Frase X”. A edição da frase é feita na aplicação “Teclado”, podendo no fim ser guardada. A utilização da frase produz um som de alerta e a frase é mostrada no ecrã.

4.3. MULTIMÉDIA

A aplicação Multimédia tem uma função lúdica para o utilizador. O utilizador pode ouvir música, ver fotografias/imagens ou assistir a filmes/vídeos. Todas estas subaplicações apresentam interface semelhante, têm botões de navegação (cima/baixo ou esquerda/direita) e no caso dos *media players* existem botões de aumento e diminuição de volume, *Play*, *Pause* e *Stop*. Nas Figura 42, Figura 43 e Figura 44 estão apresentados os ambientes gráficos do reprodutor de música, reprodutor de vídeo e visualização de imagem, respectivamente.

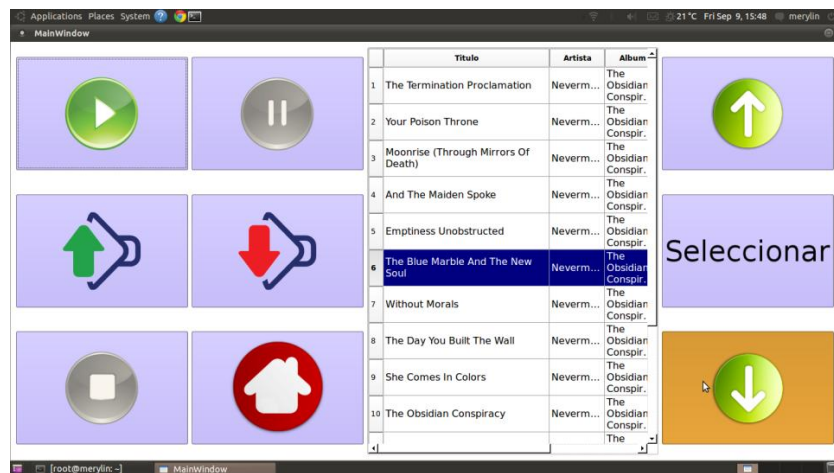


Figura 42. Reprodutor de música.

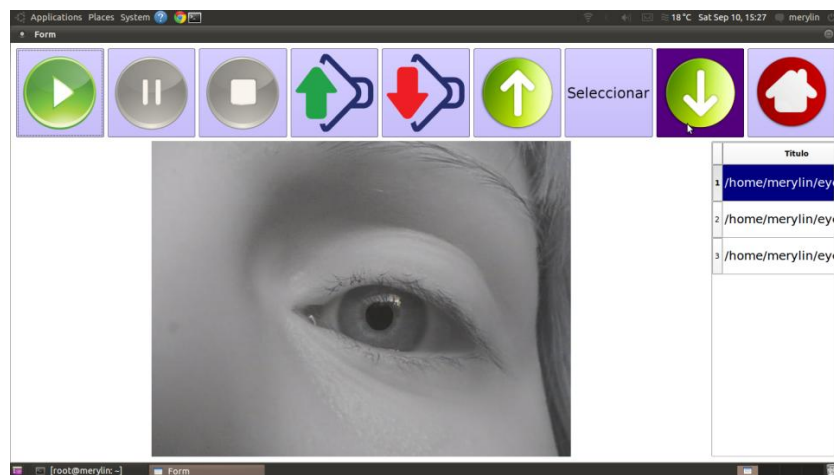


Figura 43. Reprodutor de vídeo.

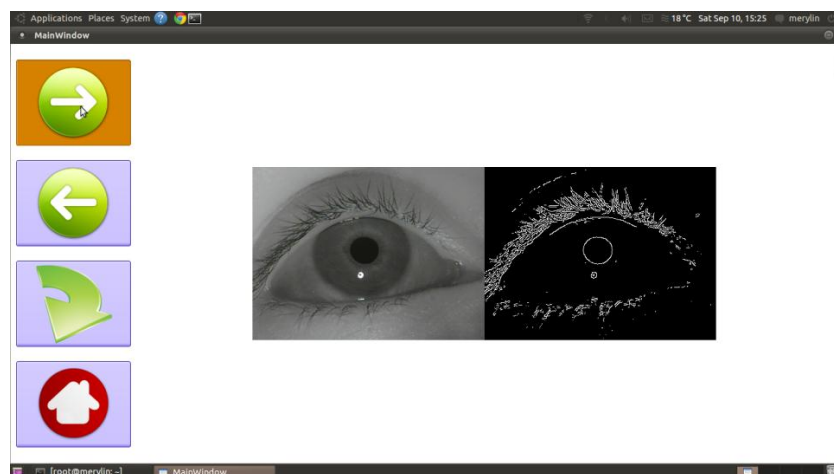


Figura 44. Visualização de imagens.

4.4. INTERNET BROWSER

Na aplicação Internet é possível navegar na internet, sendo esta uma aplicação de entretenimento/informação. A Figura 45 mostra o ambiente gráfico desta aplicação.

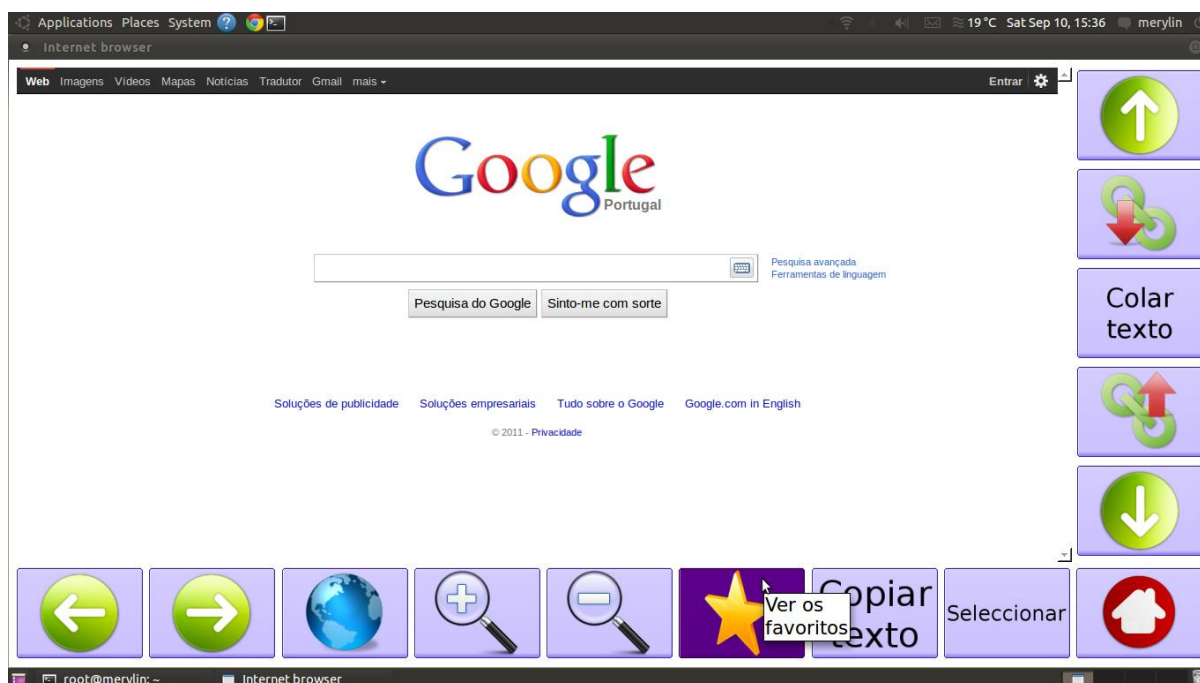


Figura 45. Internet *browser*.

Esta aplicação é semelhante a um *Web browser* com algumas alterações que visam uma navegação sem limitações. Como o sistema não tem resolução suficiente não é possível ao utilizador clicar nas hiperligações mais pequenas. Contudo, estas podem ser acedidas através dos botões com os elos. Ao serem clicados, a hiperligação seguinte (ou anterior) é focada. A hiperligação pode ser acedida ao clicar o botão "Seleccionar".

Para todas as acções que impliquem a escrita, o utilizador deve clicar em "Copiar texto", abrindo assim a aplicação do "Teclado" que permite copiar o texto. Ao voltar ao *browser*, o utilizador pode colar o texto através do botão "Colar texto".

A inexistência de uma barra de endereço deve-se, mais uma vez, à resolução do sistema. Para aceder a uma nova página da internet, o utilizador deve clicar no botão do globo. Isto abre a aplicação do "Teclado" permitindo escrever o endereço da internet.

O tamanho de todos os botões de navegação adequa-se à resolução do sistema.

4.5. CONFIGURAÇÕES

Esta é a aplicação de início do programa, mas pode ser acedida sempre que o *eye tracking* não esteja a responder como desejado. Para a calibração, o utilizador precisa sempre de ajuda, pois o cursor deixa de ser movido pelos olhos, passando a ser movido pelo rato. Por isso é mostrada uma mensagem de alerta no caso de já existir uma calibração anterior.

A configuração é de interface simples, reduzindo todos os obstáculos que o utilizador possa encontrar. Na Figura 46 é apresentada a interface gráfica.

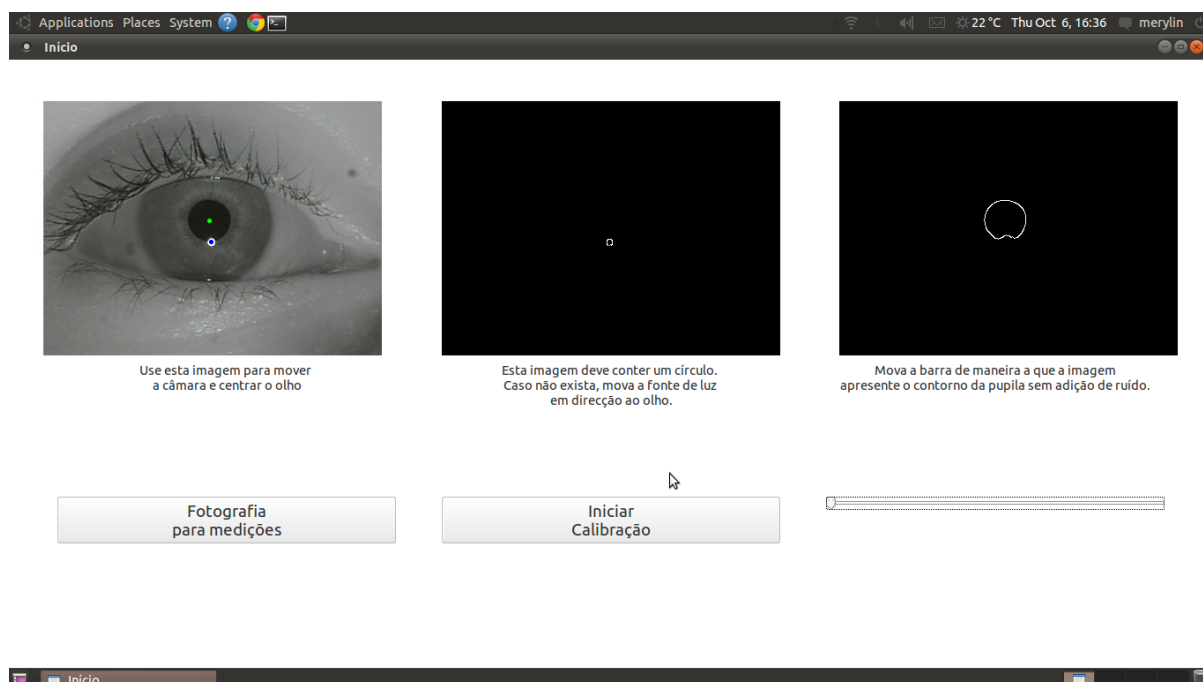


Figura 46. Configuração

Existem três imagens nesta aplicação. A primeira mostra a imagem capturada pela câmara para assim ajudar o utilizador a mover a câmara para centrar o olho. A segunda mostra a detecção do *glint*, que indica se a fonte de luz está ligada ao apresentar o contorno do *glint*. A terceira imagem mostra a detecção da pupila.

Existe ainda uma barra deslizante, que altera um parâmetro do operador *canny* para a detecção da pupila. O utilizador deve fazer deslizar a barra até a terceira imagem fornecer a informação correcta, isto é, até os contornos da pupila serem perceptíveis sem a existência de ruído.

Quando são detectados o *glint* e a pupila, aparecem dois pontos na primeira imagem, a indicar o centro de cada círculo. Se estes dois pontos não forem estáveis, é necessário redefinir o

diâmetro da pupila, através do botão “Fotografia para medições” e/ou afastar ou aproximar a fonte de luz IV.

Ao premir o botão “Fotografia para medições”, é parada a captura de imagem para ser medida a pupila. A medição da pupila é feita com o botão direito do rato pressionando-o desde uma extremidade da pupila e libertando-o na outra, traçando assim o diâmetro.

Quando a parametrização estiver concluída, o utilizador deve começar a calibração clicando no botão “Iniciar calibração”. Neste instante o utilizador deve estar atento aos pontos que lhe são mostrados sucessivamente no ecrã e fixar o seu olhar neles. A Figura 47 mostra todos os pontos de calibração.

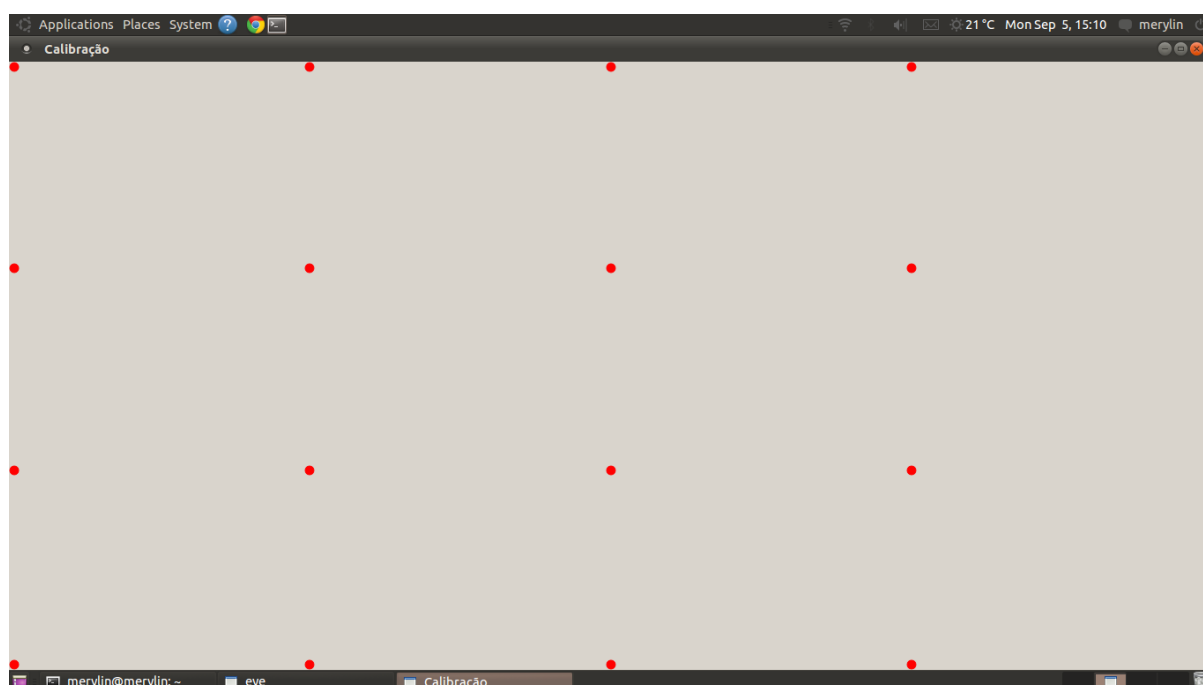


Figura 47. Pontos de calibração.

No início, os pontos têm uma cor azul que depois passa a vermelho, significando esta mudança que está a ser feita a recolha dos dados. A fase em que o ponto apresenta cor azul é uma fase de transição para o utilizador passar de um ponto para o outro e pestanejar.

Cada ponto permanece 1 segundo no estado azul e 3,5 segundos no estado vermelho. Assim, a calibração demora 1 minuto e 30 segundos. O tempo da configuração varia com o utilizador e condições ambiente, mas em média, deve demorar 2 minutos, fazendo um tempo total de 3 minutos e 30 segundos.

4.6. CLIQUE

A condição física do utilizador limita-o aos movimentos do olho e ao piscar do mesmo, logo o clique tem de ser contemplado no projecto. Existem dois métodos para realizar o clique, através de um temporizador ou através do piscar do olho.

O método do temporizador consiste na temporização da posição do cursor. Se o cursor se manter numa determinada posição (\pm *offset*) durante um tempo estabelecido, é simulado o clique do rato. Este método foi testado, e apesar do seu bom funcionamento, não mostrou ser a opção mais conveniente para o utilizador. Isto porque o utilizador podia não se aperceber do tempo decorrido e clicar onde não era intenção.

O método do piscar do olho é o actualmente usado. O utilizador pisca o olho e o clique do rato é simulado. No entanto, este piscar do olho não pode ser confundido com o piscar frequente e natural. Para isso, o piscar do olho deve ser um pouco mais prolongado, por volta de 1 segundo. A falta do *glint* é vista pelo software como o piscar do olho.

A simulação do clique do rato foi dificultada pelo facto de não haver maneira de o fazer através do sistema operativo. A solução encontrada foi saber em qual dos botões estava o rato, aquando o emitir do sinal de clique (como descrito no fluxograma).

5. ESTRUTURA DE APOIO

A estrutura de apoio é um componente não obrigatório para a utilização do sistema. Contudo, é importante apresentar uma solução para o fixar e manusear. Foi utilizado o *SolidWorks 2011* para desenhar a estrutura em 3D.

A solução encontrada é baseada num design simples, semelhante às mesas de apoio a refeições na cama. É constituída por duas mesas de apoio, podendo o computador ser colocado na mesa inferior e o ecrã com a câmara na mesa superior, como visível na Figura 48.

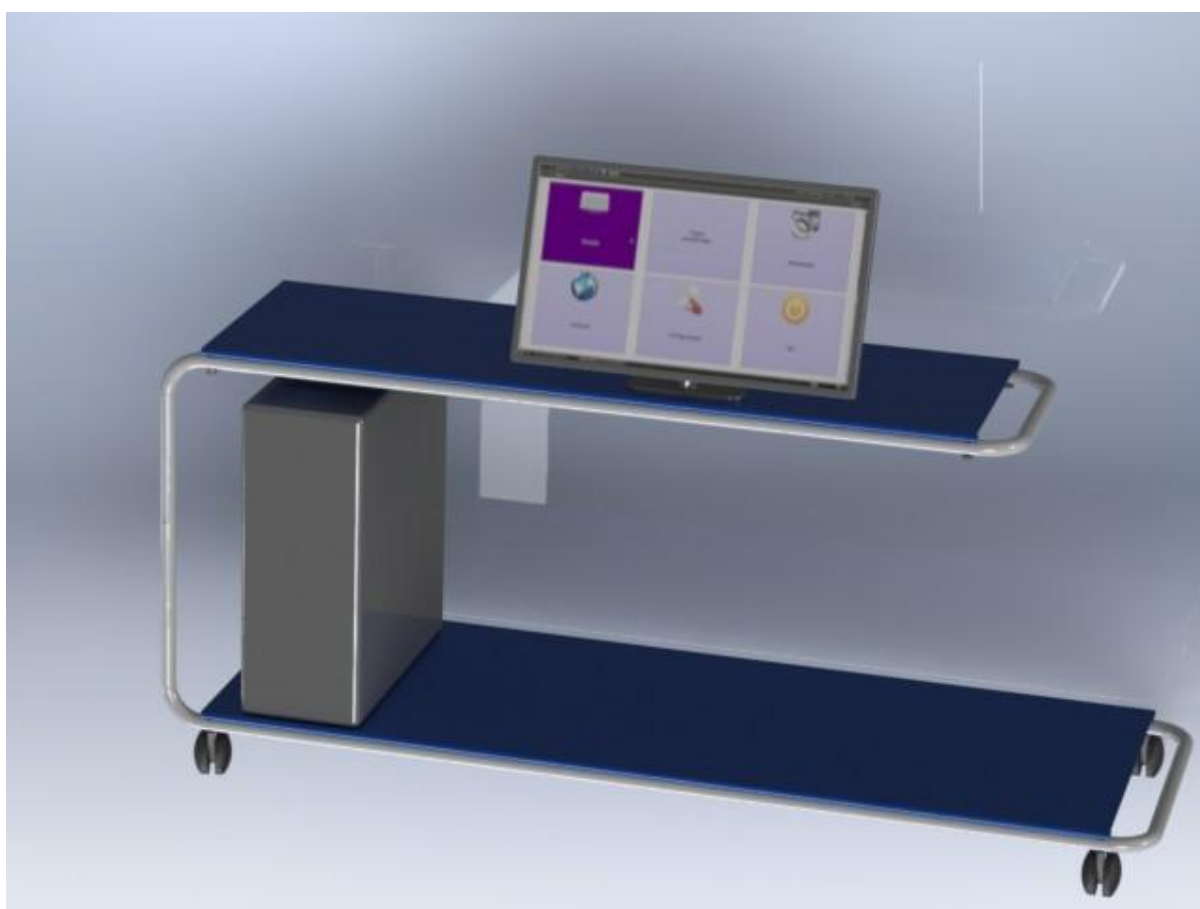


Figura 48. Estrutura de apoio.

A distância entre a mesa superior e a inferior é regulável através dos orifícios existentes no tubo que une os dois níveis de apoio (Figura 49), podendo assim adaptar-se à altura de uma cama ou de uma cadeira. A altura mínima é de 70 cm e a altura máxima é de 110 cm.

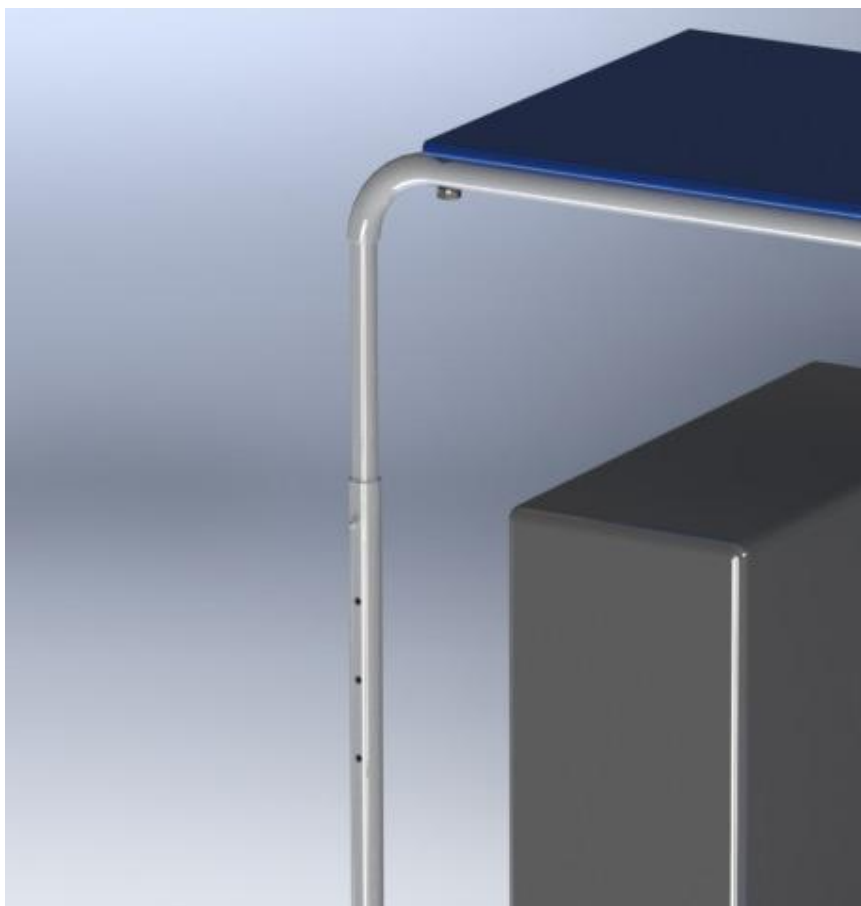


Figura 49. Mecanismo de selecção de alturas.

A estrutura apresenta um peso de 9 kg já com ecrã e computador, e está suportada sobre 4 rodas, permitindo assim o deslocamento do sistema.

6. CONCLUSÃO

No âmbito desta tese foi desenvolvido um sistema de *eye tracking* como interface humano-máquina para humanos com incapacidade motora. Foi também desenvolvida uma interface gráfica de acordo com as necessidades do utilizador, assim como uma estrutura de apoio adaptável em altura, para que se possa colocar o sistema ao nível do utilizador.

O desenvolvimento do *eye tracking* foi a fase com mais obstáculos causados, na sua maioria, à escolha inicial do hardware. A selecção da câmara foi limitada no tempo, devido à necessidade de se iniciar o projecto. Desta forma, a câmara escolhida mostrou ter uma resolução insuficiente para a utilização que se pretendia desta.

Como não foi possível entrar em contacto com futuros utilizadores, a realização da interface gráfica durou mais tempo que o esperado, isto porque surgiram questões sobre as dimensões, características e funções em todas as aplicações. Assim, a interface gráfica foi muito apoiada nas soluções existentes, até porque estas foram desenvolvidas em conjunto com médicos especialistas e futuros utilizadores.

A estrutura de apoio apresenta um design simples e compreende todas as características propostas no início do projecto sendo adaptável em altura, de maneira a se colocar ao lado de uma cama ou uma cadeira. A estrutura concebida é apenas uma sugestão, podendo não ser necessária na maioria dos casos.

O sistema final encontrado não permite variações de distância entre o utilizador, a câmara e o ecrã. O utilizador deve estar a uma distância de 60 cm do ecrã e a uma distância de 30 cm da câmara, aproximadamente. O uso do sistema com outras configurações não terá um comportamento previsível, isto porque se o ecrã e/ou a câmara estiverem mais afastados do utilizador a variação do movimento do olho entre dois pontos do ecrã é menor, diminuindo assim a resolução. Numa situação inversa, onde a câmara está mais próxima (ou é de maior resolução) ou o ecrã está mais próximo (ou é de maior dimensão) é provável que o sistema se torne mais estável ou com maior resolução, mas este comportamento também não foi estudado.

O uso do sistema está também limitado a utilizadores sem problemas na visão pois não foi testado com utilizadores com pálpebras caídas, pupilas assimétricas, olhos secos, nistagmo ou

estrabismo. Foi testado com pessoas com óculos e lentes de contacto sem terem sido detectados problemas no seu funcionamento. Contudo, não foi testado com pessoas de olhos claros, mas é previsível que o comportamento do sistema não se altere, pois uma íris clara contribui para um maior contraste entre a íris e a pupila.

Em conclusão, os objectivos inicialmente propostos foram atingidos. Foi produzido um sistema com interface humano-máquina através dos olhos, com uma aplicação para a comunicação escrita. É também um sistema de custo reduzido e aberto, no sentido em que o hardware pode ser mudado sem mudar radicalmente as características originais, isto é, pode-se usar outra câmara, ecrã ou computador, desde que não difiram muito do utilizado. Quanto ao software, apesar de o computador correr obrigatoriamente o sistema operativo Linux na distribuição *Ubuntu* versão 10.04, este é gratuito.

6.1. CARACTERÍSTICAS DO SISTEMA

Este sistema é direccionado para a tecnologia de assistência. É composto por uma câmara sensível aos raios infravermelhos, um computador com sistema operativo Linux, um ecrã com, no mínimo, 15,6 polegadas e uma fonte de luz de infravermelhos. Existe também a possibilidade de integrar uma estrutura de apoio.

Para o *eye tracking* é usado o método *dark pupil* a uma frequência de 25 Hz. São permitidos movimentos ligeiros (2 a 3 cm) apenas para os lados e na vertical. Os movimentos em profundidade não são contemplados neste sistema. A calibração do sistema leva aproximadamente 1,5 minutos a ser realizada.

As aplicações existentes são, o Teclado, Frases predefinidas, Multimédia e Internet browser. O clique, para a interacção nas aplicações, é realizado pelo piscar do olho.

6.2. PROPOSTAS PARA TRABALHO FUTURO

Propõe-se que no futuro seja seleccionado outro equipamento para a captura de vídeo, nomeadamente, uma câmara com maior resolução, podendo ser necessárias modificações no software de *eye tracking*. Seria útil também estudar quais as vantagens em usar os dois métodos (*bright* e *dark pupil*) em simultâneo na realização do *eye tracking*.

Apesar de a comunicação ser um dos principais obstáculos de pessoas com esclerose lateral amiotrófica, existem outro tipo de acções que permitiriam aos utilizadores maior independência. Fala-se aqui do accionamento de luzes, portas, cadeiras de rodas, televisões, etc.

É importante também referir que existem aplicações que à primeira vista podem não ser valorizadas mas, por exemplo, através de jogos, o utilizador exercita a mente, atingindo assim outro nível de saúde mental. É então proposto o estudo e desenvolvimento de aplicações com este fim.

Por falta de tempo, não se desenvolveu uma aplicação capaz de enviar mensagens escritas através de um cartão SIM ou através da internet. Esta seria uma aplicação que poderia ser usada, por exemplo, numa situação de emergência em que o utilizador se encontrasse desacompanhado.

7. REFERÊNCIAS

- [1] Ivan E. Sutherland, "Sketchpad: A man-machine graphical communication system," in Proceedings of the Spring Joint Computer Conference, 1963, pp. 329-346.
- [2] Hewett et al., ACM SIGCHI Curricula for Human-Computer Interaction., 2009.
- [3] Andrew Duchowski, Eye Tracking Methodology - Theory and Practice - Second Edition.: Springer, 2007.
- [4] C.H. Morimoto, D. Koons, A. Amir, and M. Flickner, "Frame-Rate Pupil Detector and Gaze Tracker," Julho 1999.
- [5] C.H. Morimoto, D. Koonsb, A. Amir, and M. Flicknerb, "Pupil detection and tracking using multiple light sources," Julho 1999.
- [6] LC Technologies, Inc. (2010, Novembro) LC Technologies Inc, Eyegaze Systems. [Online]. <http://www.eyegaze.com/>
- [7] COGAIN. (2010, Dec.) COGAIN: Communication by Gaze Interaction. [Online]. http://www.cogain.org/wiki/Main_Page
- [8] EyeTech Digital Systems. (2010, Dezembro) EyeTech Digital Systems - Eye tracking technology solutions. [Online]. <http://www.eyetechds.com/>
- [9] tobii. (2010, Dezembro) Tobii Technology - eye tracking and eye control | usability testing, market research and scientific research | AAC communication solutions. [Online]. <http://www.tobii.com/>
- [10] DynaVox. (2010, Dec.) Communication Devices - Speech Devices | DynaVox. [Online]. <http://uk.dynavoxtech.com/>
- [11] alea technologies. (2010, Dec.) alea technologies. [Online]. <http://www.alea-technologies.com/>
- [12] Humanelektronik GmbH · Adaptive & Assistive Technologies. (2010, Dec.) SEETECH. [Online].

<http://www.see-tech.de/>

[13] Prentke Romich Company. (2010, Dec.) Global Leader in the Augmentative and Alternative Industry. [Online]. <http://www.prentrom.com/>

[14] The EyeWriter. (2010, Dezembro) EyeWriter. [Online]. <http://www.eyewriter.org/>

[15] PlayStation. (2011, Janeiro) PlayStation® Eye for the PS3™. [Online].
<http://us.playstation.com/ps3/accessories/playstation-eye-camera-ps3.html>

[16] Peau Productions. (2011, Janeiro) Peau Productions. [Online].
<http://peauproductions.com/cameras.html>

[17] OpenCV. (2011, Janeiro) OpenCV Wiki. [Online]. <http://opencv.willowgarage.com/wiki/>

APÊNDICE I

Percorre-se a imagem com o objectivo de rotular os pixéis de acordo com a sua vizinhança, através de um padrão (*kernel*) apresentado na Figura 50.

A	B	C
D	X	

Figura 50. *Kernel* de rotulagem.

É necessário um buffer com o mesmo tamanho da imagem para guardar os rótulos de todos os pixéis e é inicializado a zero. Começando pelo topo do lado esquerdo e acabando em baixo no lado direito da imagem, é aplicado o *kernel* a todos os pixéis na posição X, onde são verificados os pixéis vizinhos A, B, C e D. O pixel X é rotulado com o rótulo mais baixo do conjunto, excepto o 0. Veja-se o exemplo da Figura 51. O algoritmo está descrito na Figura 52.

			A	B	C
			D	X	

0	0	0	0	0	0
0	1	0	0	0	0
0	1	1	1	0	0
0	0	0	1	0	2
0	0	0	0	X	0
0	0	0	0	0	0

Figura 51. Aplicação do *kernel*.

```

label = 1; // Label inicial
Para cada pixel da imagem
{
    Se o pixel X é preto
    {
        Se A, B, C e D são iguais a zero
        {
            Rotular pixel X com label;
            Incrementar label;
        }
        Se não
        {
            Rot = Menor valor de A, B, C e D excepto 0;
            Rotular pixel X com Rot;
        }
    }
}

```

Figura 52. Algoritmo de detecção de *blobs*.

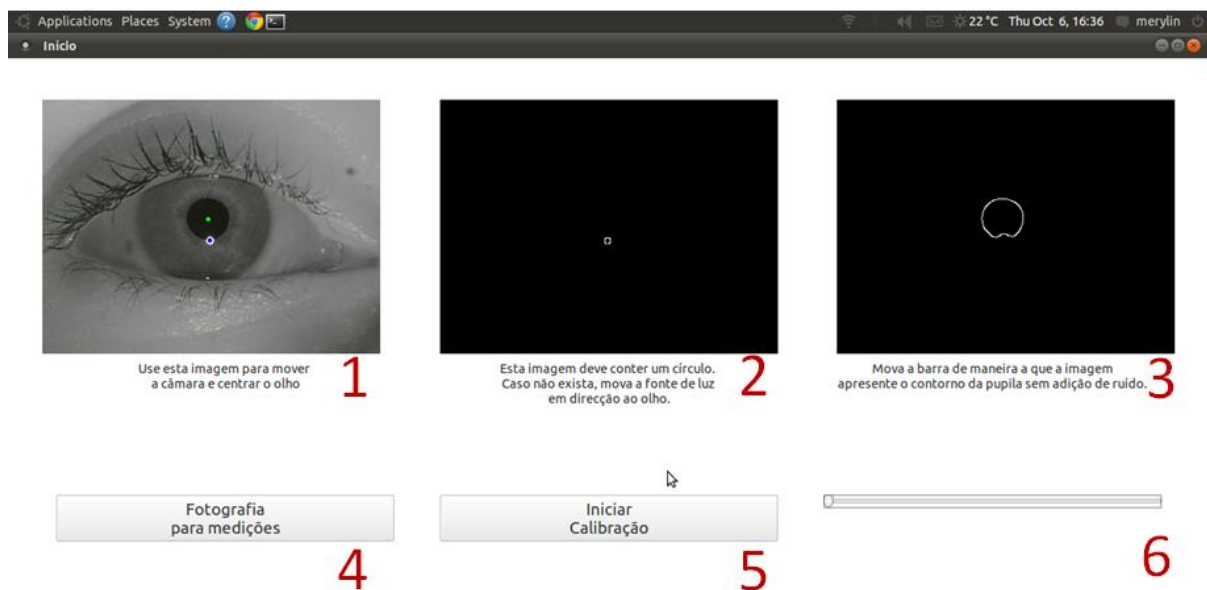
APÊNDICE II

Antes de utilizar

No início, o ficheiro executável “EyeTracking” deve ser copiado para o computador. Na mesma directoria devem ser criadas 3 pastas, nomeadamente, “Música”, “Vídeo” e “Imagem”. Posteriormente, devem ser colocados os ficheiros correspondentes a cada pasta. Se correctamente colocados, os ficheiros depois são visualizados pelo utilizador. Deve ser criada também outra pasta, “Texto”. Ao longo da utilização do programa, os ficheiros de texto que o utilizador decidir guardar, são colocados nesta pasta.

Antes de iniciar o programa, a câmara deve estar ligada ao computador, e os *leds* ligados à alimentação. Se por alguma razão um dos componentes não funcionar, terão de ser substituídos antes de se iniciar o programa.

Iniciar o programa



- 1) Imagem capturada pela câmara.
- 2) Detecção do *glint*.
- 3) Detecção da pupila.
- 4) Fotografia para medições.
- 5) Iniciar calibração.
- 6) Barra deslizante.

O programa é sempre iniciado com a configuração do sistema.

Existem três imagens nesta aplicação. 1) mostra a imagem capturada pela câmara para assim ajudar o utilizador a mover a câmara para centrar o olho. 2) mostra a detecção do *glint*, que indica se a fonte de luz está ligada ao apresentar o contorno do *glint*. 3) imagem mostra a detecção da pupila.

O utilizador deve fazer deslizar 6) até 3) fornecer a informação correcta, isto é, até os contornos da pupila serem perceptíveis sem a existência de ruído.

Uma boa configuração produz dois pontos em 1). Um ponto verde, a indicar o centro da pupila, e um ponto azul, a indicar o centro do *glint*. Se os pontos se apresentarem inconstantes, o diâmetro da pupila deve ser redefinido e/ou a fonte de luz deve ser aproximada ou afastada do utilizador até os pontos estabilizarem. Para redefinir o diâmetro da pupila, deve ser clicado 4) e com o botão direito do rato traçar o diâmetro da pupila. Outra medida para estabilizar os pontos, é voltar a mover 6) de maneira a obter uma circunferência mais definida.

Assim que os pontos estabilizem, a calibração deve ser iniciada.

Calibração



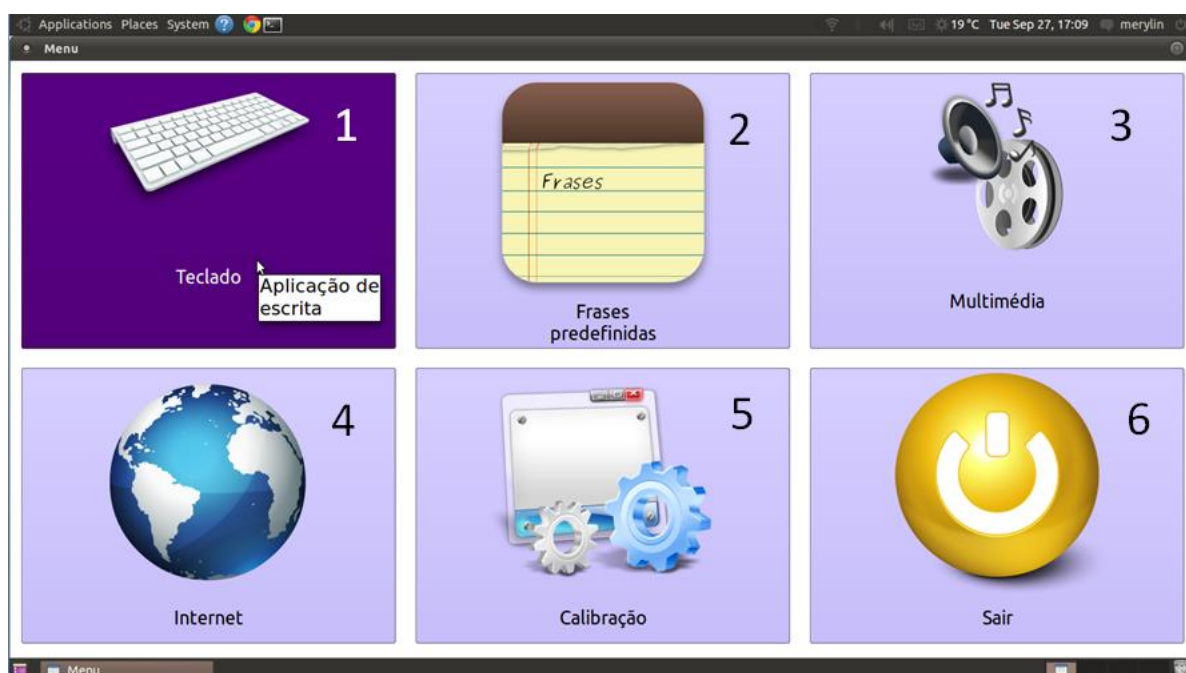
Durante esta fase, são apresentados pontos no ecrã. Estes pontos devem ser fixados com o olhar. Os pontos apresentam duas cores, azul e vermelho. Quando o ponto está vermelho, o utilizador deve evitar pestanejar e desviar o olhar. Enquanto o ponto está azul, o utilizador deve relaxar, pestanejar e olhar para o próximo ponto.

Uma má calibração pode inutilizar o programa. Desta forma, a calibração deve ser feita com calma e atenção. Se o programa não corresponder ao esperado, é necessária uma nova calibração.

Interface

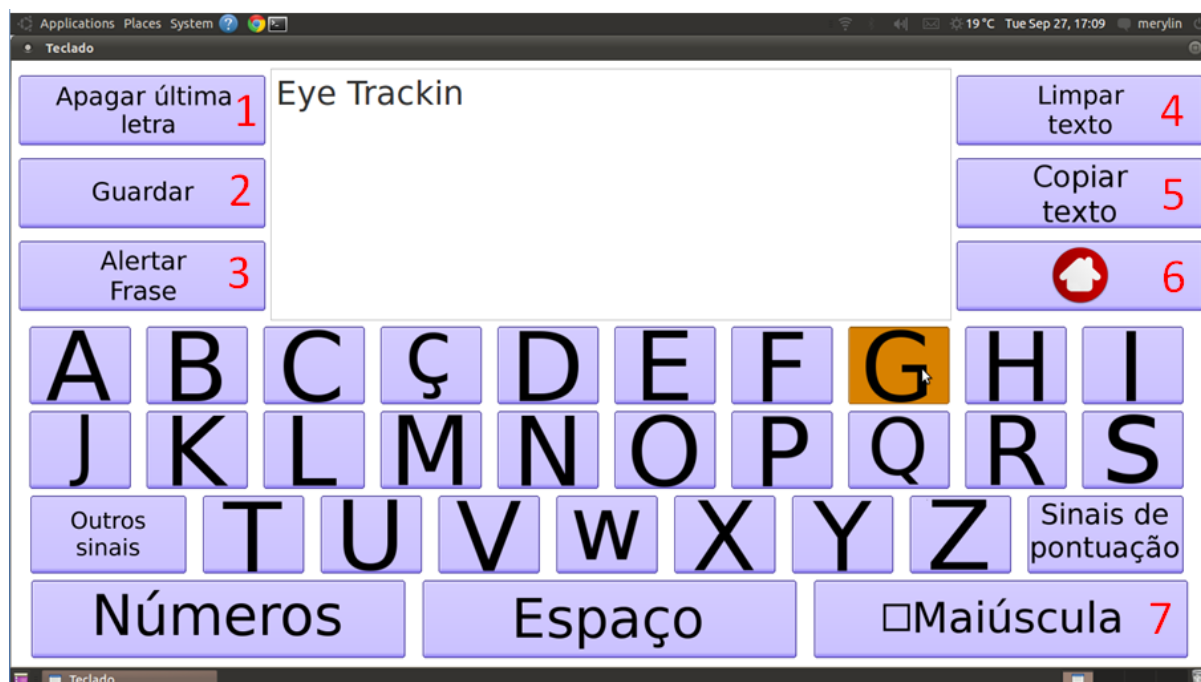
Para o clique dos botões, o utilizador deve piscar o olho durante 1 segundo, aproximadamente.

Menu inicial



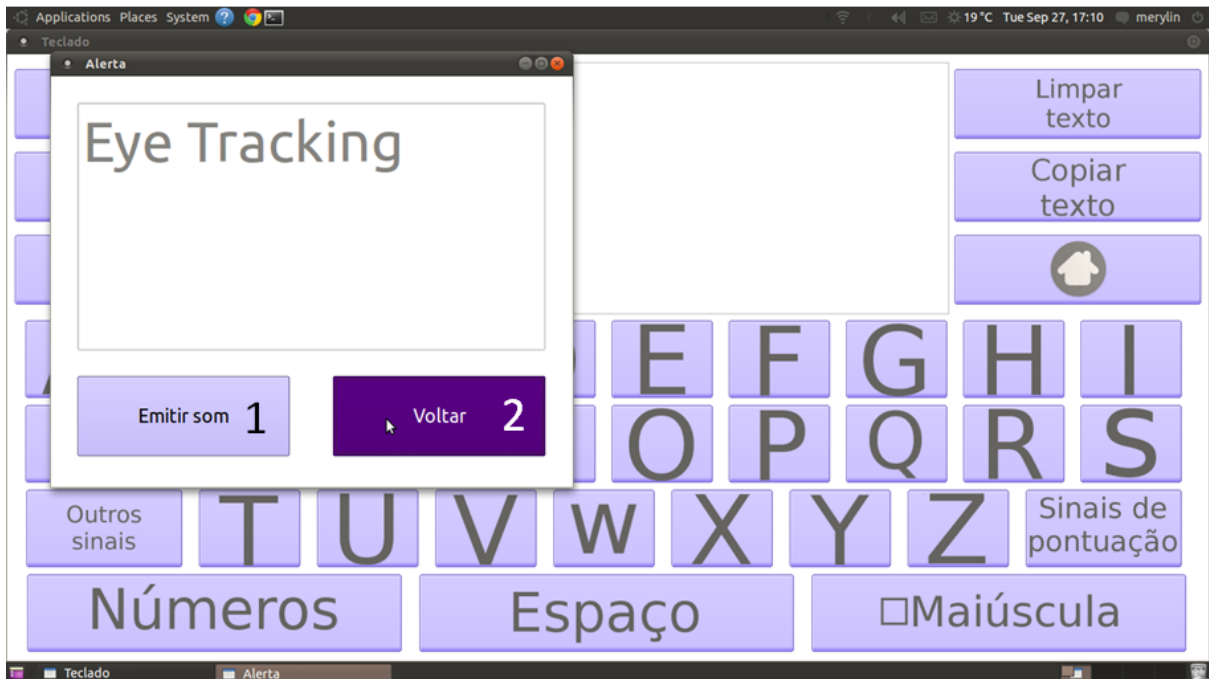
- 1) Aplicação Teclado - Escrita
- 2) Aplicação Frases Predefinidas - Frases rápidas
- 3) Aplicação Multimédia - Ouvir música, ver imagens e vídeos
- 4) Aplicação Internet - Navegar na internet
- 5) Aplicação Calibração - Recalibrar o sistema
- 6) Sair do sistema

Teclado



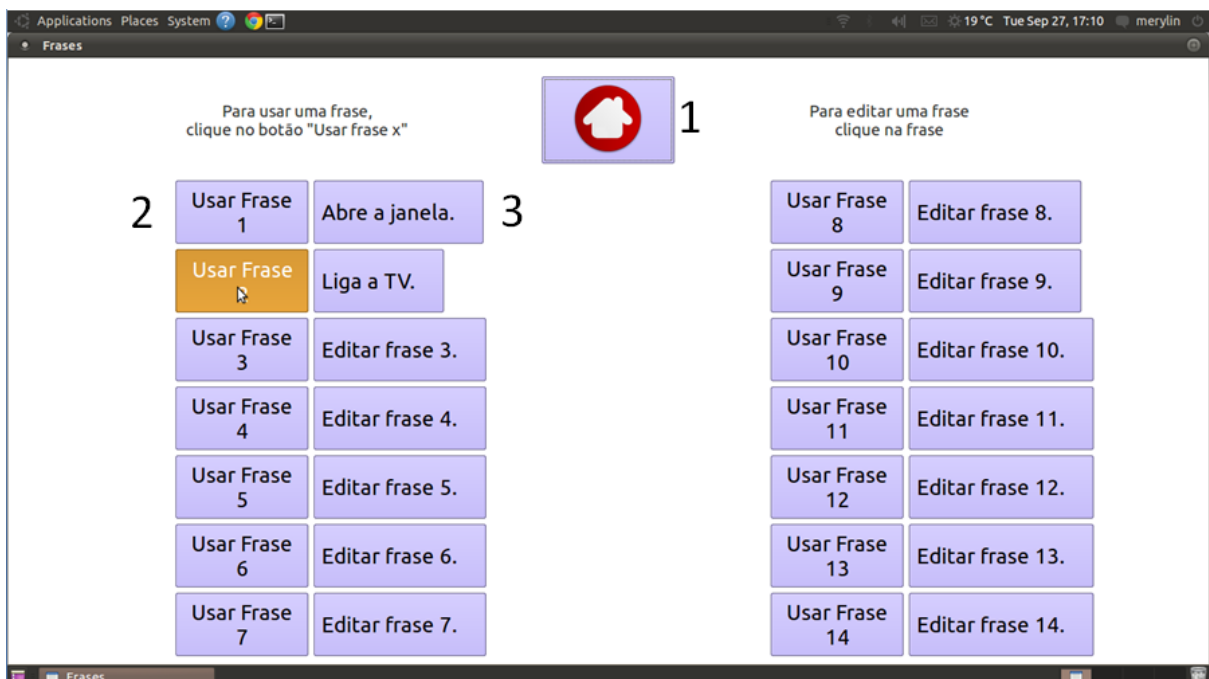
- 1) Apagar a última letra escrita
- 2) Guarda o texto presente
- 3) Apresenta alerta com som
- 4) Apaga o texto presente
- 5) Copia o texto presente
- 6) Volta ao Menu inicial
- 7) Quando seleccionado apresenta um marcador

Alerta



- 1) Volta a emitir som de alerta
- 2) Volta à aplicação Teclado

Frases predefinidas



- 1) Volta ao Menu inicial
- 2) Apresenta alerta com som
- 3) Edita a frase na aplicação Teclado

Multimédia



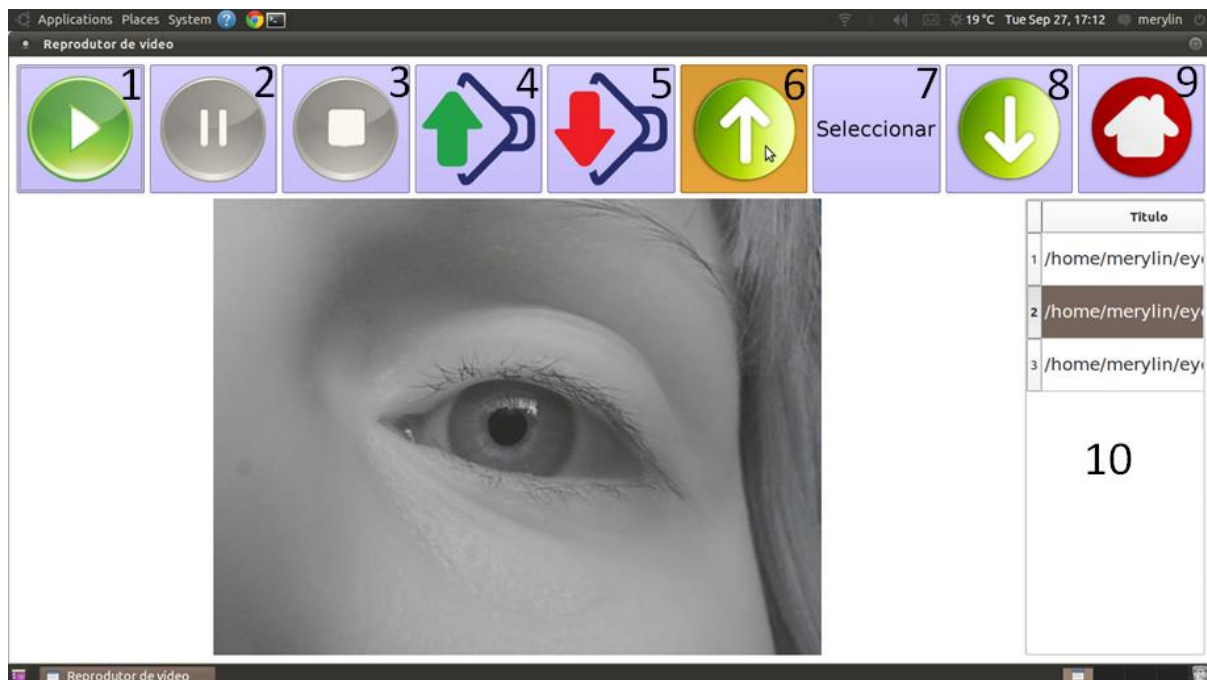
- 1) Aplicação Reprodutor de música
- 2) Aplicação Reprodutor de vídeo
- 3) Aplicação Visualização de imagens
- 4) Volta ao Menu inicial

Reprodutor de música



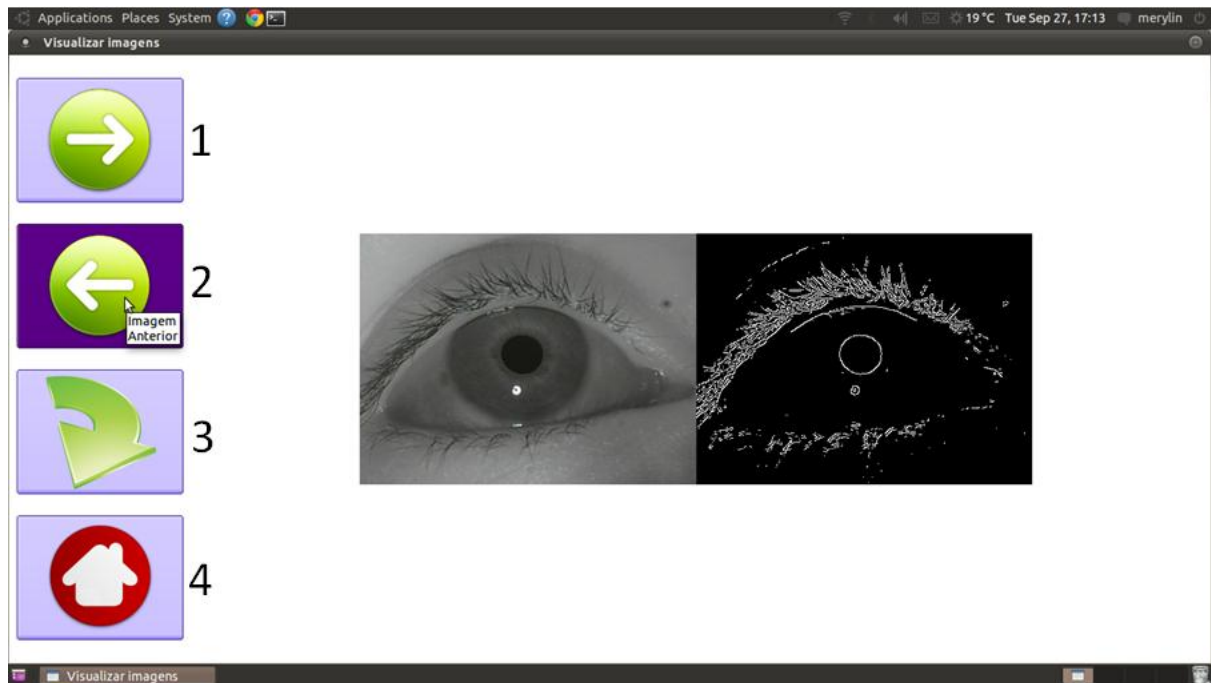
- 1) Reproduzir música
- 2) Pausar música
- 3) Aumentar o volume
- 4) Diminuir o volume
- 5) Parar a reprodução de música
- 6) Voltar ao Menu inicial
- 7) Navegar para cima na tabela de conteúdos
- 8) Seleccionar item da tabela de conteúdos
- 9) Navegar para baixo na tabela de conteúdos
- 10) Tabela de conteúdos

Reprodutor de vídeo



- 1) Reproduzir vídeo
- 2) Pausar reprodução de vídeo
- 3) Parar reprodução de vídeo
- 4) Aumentar o volume
- 5) Diminuir o volume
- 6) Navegar para cima na tabela de conteúdos
- 7) Seleccionar item da tabela de conteúdos
- 8) Navegar para baixo na tabela de conteúdos
- 9) Voltar ao Menu inicial
- 10) Tabela de conteúdos

Visualização de imagens



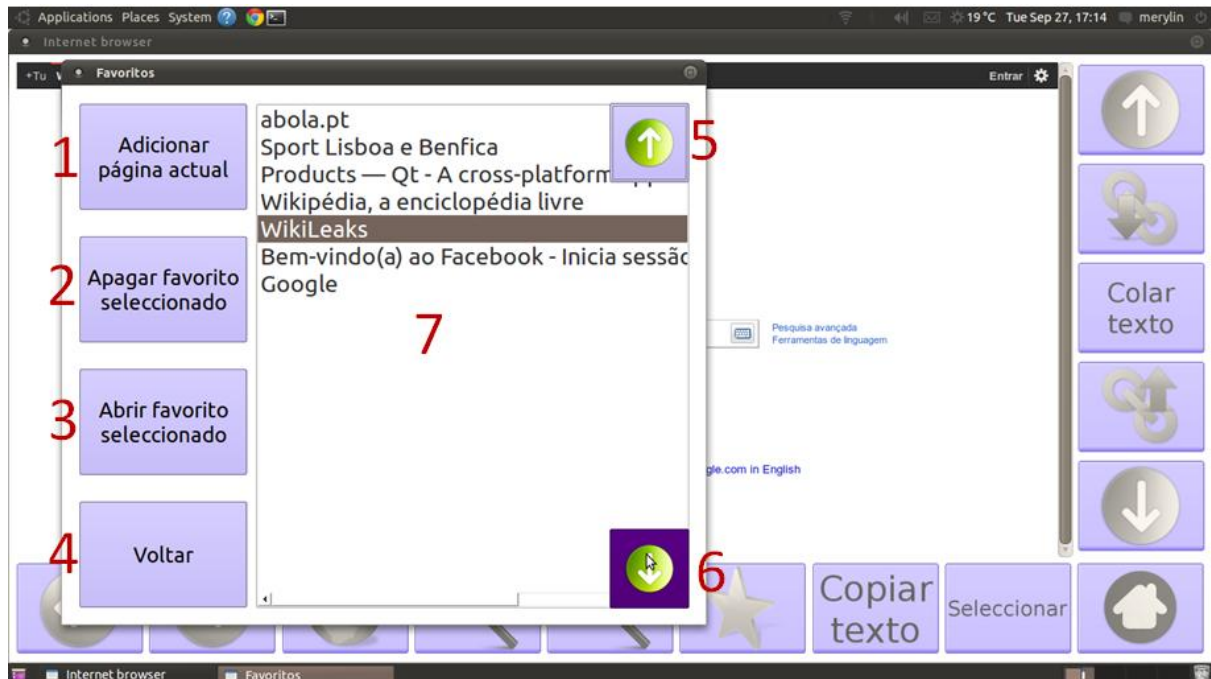
- 1) Navegar para a imagem seguinte
- 2) Navegar para a imagem anterior
- 3) Rodar a imagem 90° para a direita
- 4) Voltar ao Menu inicial

Internet



- 1) Voltar à página anterior
- 2) Voltar à página seguinte
- 3) Inserir endereço de internet
- 4) Aumentar ampliação da página
- 5) Diminuir ampliação da página
- 6) Aplicação Favoritos
- 7) Escrever texto na aplicação Teclado
- 8) Seleccionar hiperligação
- 9) Voltar ao Menu inicial
- 10) Navegar para baixo na página
- 11) Navegar para a hiperligação anterior
- 12) Colar texto
- 13) Navegar para a hiperligação seguinte
- 14) Navegar para cima da página

Favoritos



- 1) Adiciona página actual à lista de favoritos
- 2) Apaga item seleccionado da lista de favoritos
- 3) Abre favorito seleccionado
- 4) Volta à aplicação Internet
- 5) Navega para cima na lista de favoritos
- 6) Navega para baixo na lista de favoritos
- 7) Lista de favoritos